

Db 1381 QMRKLSQSWDRCKDGLYALKETHPHILMELMEKSKKGLPLSS 1429

RESULT 2
ABG97969
ID ABG97969 standard; Protein; 1429 AA.
XX ABG97969;
XX 07-JAN-2003 (first entry)
DE Human leucine rich repeat domain containing protein #2.
XX
XX Leucine rich repeat; nervous system; human; neural disorder; apoptosis;
XX renal disorder; immune disorder; arthritis; asthma; AIDS;
XX acquired immunodeficiency syndrome; rheumatoid arthritis;
XX hematopoietic disorder; metabolic disorder; reproductive disorder;
XX pulmonary disease; cardiovascular disease; hyperproliferative disorder;
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX Huntington's disease; developmental disorder; autoimmune disease;
XX Addison's disease; hemolytic anaemia; antiphospholipid syndrome;
XX allergic encephalomyelitis; gene therapy.
XX
XX Homo sapiens.
XX
XX WC00274959-A2.
XX
XX 26-SEP-2002.
XX
XX 20-DEC-2001; 2001WC-US50457.
XX
XX 03-JAN-2001; 2001US-259479P.
XX
XX 09-JAN-2001; 2001US-260616P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Ramanathan C, Feder J, Mintier G;
XX
XX WPI; 2002-750554/81.
XX
XX New HLRNSI nucleic acids and polypeptides, useful for preventing,
XX treating, or ameliorating e.g. renal disorder, immune, hematopoietic,
XX metabolic, reproductive, pulmonary, cardiovascular or autoimmune
XX diseases
XX
XX Example 1; Page 392-397; 415pp; English.
XX
XX The invention describes nucleic acids encoding human leucine-rich repeat
XX containing proteins expressed in nervous system tissues, HLRNSI. The
XX HLRNSI polypeptide or the polynucleotide is useful for preventing,
XX treating, or ameliorating a neural disorder or a disorder related to
XX aberrant apoptosis modulation (either directly or indirectly), renal
XX disorder, immune disorder (e.g. arthritis, asthma, acquired
XX immunodeficiency syndrome (AIDS) or rheumatoid arthritis),
XX hematopoietic, metabolic, reproductive, pulmonary or cardiovascular
XX diseases, hyperproliferative disorders, neurodegenerative diseases
XX (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
XX disease), developmental disorders, non-infectious disorders, nervous
XX system diseases and/or disorders, and autoimmune diseases (e.g.
XX Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or
XX allergic encephalomyelitis). The polynucleotides are also useful as
XX chromosome markers for chromosome identification, gene therapy, and in
XX identifying organisms from minute biological samples. This is the amino
XX acid sequence of a leucine-rich repeat containing protein.
XX
XX Sequence 1429 AA;
XX
XX Query Match 100.0%; Score 7534; DB 23; Length 1429;
XX Best Local Similarity 100.0%; Pired. No. 0;
XX Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MAGGAMRLACYLEFLKKEELKEFOLLANKHSSSSSGETPAQPEKTSQMEVASYLVAQ 60

QY 61 YGSRAMDLALHTWEMGLRSLCAQAOEGAGHSPSPFYSSEPHLSPSPQSTPAVAMPV 120

Db 61 YGSRAMDLALHTWEMGLRSLCAQAOEGAGHSPSPFYSSEPHLSPSPQSTPAVAMPV 120

QY 121 IHELPACTGSGERRVLRQLPDTSGRRMRREISASLLYQALPSSPDHSPQESNAPTST 180

Db 121 IHELPACTGSGERRVLRQLPDTSGRRMRREISASLLYQALPSSPDHSPQESNAPTST 180

QY 181 AVTGSWSPPOSLAPRECAPPTQWPLDFTSGIYYTEIRERREKSEKRPMAAVVGT 240

Db 181 AVTGSWSPPOSLAPRECAPPTQWPLDFTSGIYYTEIRERREKSEKRPMAAVVGT 240

QY 241 PPOAHNLSOPHHHPWBSVRESICSTWPKNENEPNOKFTOLLQRPHPSSODPLVRWS 300

Db 241 PPOAHNLSOPHHHPWBSVRESICSTWPKNENEPNOKFTOLLQRPHPSSODPLVRWS 300

QY 301 PDVVEENRGLIIRDLFGGLDTPRPRIYILQAGAGIGSTLARQYKEMGRGQLYGDR 360

Db 301 PDVVEENRGLIIRDLFGGLDTPRPRIYILQAGAGIGSTLARQYKEMGRGQLYGDR 360

QY 361 FOHVFFSCRELAQSKVSLAEILIGDGTATPAPIRQILSRPRLFLIDGVDEPGVILQ 420

Db 361 FOHVFFSCRELAQSKVSLAEILIGDGTATPAPIRQILSRPRLFLIDGVDEPGVILQ 420

QY 421 EPSESELCHWSQOPADALLIGSLIGTILPEASFLITARTTAONLIPSLQARWVVG 480

Db 421 EPSESELCHWSQOPADALLIGSLIGTILPEASFLITARTTAONLIPSLQARWVVG 480

QY 481 PSESSRKEYFYRYETBEROAIRAFRLVKSKEEMALCLVWVSWLACTCMQMKREKL 540

Db 481 PSESSRKEYFYRYETBEROAIRAFRLVKSKEEMALCLVWVSWLACTCMQMKREKL 540

QY 541 TLTSKTTTTLCLHYLAQALQAPLQPOLRDLCSLAABGIWQKTLFSPDDLRRHGLDGI 600

Db 541 TLTSKTTTTLCLHYLAQALQAPLQPOLRDLCSLAABGIWQKTLFSPDDLRRHGLDGI 600

QY 601 ISTFLKMGILQHPPIPLSYSFHLCQEPFAANSVYLEDKRGKXNCTIDLEKTEAY 660

Db 601 ISTFLKMGILQHPPIPLSYSFHLCQEPFAANSVYLEDKRGKXNCTIDLEKTEAY 660

QY 661 GIGLFGASTTRFLGLSDGEBREMENTFHCRLSGGRNLQWVPSQLLQPHSLESLH 720

Db 661 GIGLFGASTTRFLGLSDGEBREMENTFHCRLSGGRNLQWVPSQLLQPHSLESLH 720

QY 721 CLYETRNKTFELQVMAHFEEMGCVEITDMLLVCTFCIKFSRHVKQLQILBGRHRTWS 780

Db 721 CLYETRNKTFELQVMAHFEEMGCVEITDMLLVCTFCIKFSRHVKQLQILBGRHRTWS 780

QY 781 PTWVVLFRWVPTVDAYWQILFSLVLTATRNKKELDLSGNSLSHAAYSLCTTLRRPCLLE 840

Db 781 PTWVVLFRWVPTVDAYWQILFSLVLTATRNKKELDLSGNSLSHAAYSLCTTLRRPCLLE 840

QY 841 TLRLAGGGLAEDCKDLAPGLRANQTLTELDLSFNVLTDAAGHLCORLQPSCKLQRLQ 900

Db 841 TLRLAGGGLAEDCKDLAPGLRANQTLTELDLSFNVLTDAAGHLCORLQPSCKLQRLQ 900

QY 901 LVSCGLTSDCCODLASYLSAPSLKEILDQONNLDVGVALCEGLRHPAKCLIRGLDQ 960

Db 901 LVSCGLTSDCCODLASYLSAPSLKEILDQONNLDVGVALCEGLRHPAKCLIRGLDQ 960

QY 961 TTSLSDEMRQELRALBOKPOLLFSRRKPSVMTPTBGLDGEBSNSTSLKRRRLSERA 1020

Db 961 TTSLSDEMRQELRALBOKPOLLFSRRKPSVMTPTBGLDGEBSNSTSLKRRRLSERA 1020

QY 1021 ASHYAQNANLKLIDVSKIFPIAEIAESSPEVVEILLCVSPASQGLIHRKPLGTDDEFW 1080

Db 1021 ASHYAQNANLKLIDVSKIFPIAEIAESSPEVVEILLCVSPASQGLIHRKPLGTDDEFW 1080

QY 1081 GPTGPVATEVVDKELLYRVHFPVAGSYRWPNTGLCFVMEAVTVEIEFCVMDQFLGEIN 1140

Db 1081 GPTGPVATEVVDKELLYRVHFPVAGSYRWPNTGLCFVMEAVTVEIEFCVMDQFLGEIN 1140

QY 1141 PHSNVAAGPLDIDKAEPAVEAVH.PHFVALQGHVDTSLPQMAHFKKEGMLLEKPARV 1200
 CC 1141 PHSNVAAGPLDIDKAEPAVEAVH.PHFVALQGHVDTSLPQMAHFKKEGMLLEKPARV 1200
 Db 1141 PHSNVAAGPLDIDKAEPAVEAVH.PHFVALQGHVDTSLPQMAHFKKEGMLLEKPARV 1200
 QY 1201 ELHHTVLENPSPSLVLLKMHNNALRFTPVTSVLLYHRVHEEVTFFH.YILPSCSIR 1260
 CC 1201 ELHHTVLENPSPSLVLLKMHNNALRFTPVTSVLLYHRVHEEVTFFH.YILPSCSIR 1260
 Db 1201 ELHHTVLENPSPSLVLLKMHNNALRFTPVTSVLLYHRVHEEVTFFH.YILPSCSIR 1260
 QY 1261 KELELCYRSPGEBQOLFSEFVYVGHLSGIRLQVNDKDETLVWEALVKPGDIMPATTLIP 1320
 CC 1261 KELELCYRSPGEBQOLFSEFVYVGHLSGIRLQVNDKDETLVWEALVKPGDIMPATTLIP 1320
 Db 1261 KELELCYRSPGEBQOLFSEFVYVGHLSGIRLQVNDKDETLVWEALVKPGDIMPATTLIP 1320
 QY 1321 ARIAVSPDAPQLHFVDQYRQOLIARVTSVEVLDKLGQYLQSOYERVALENTRP 1380
 CC 1321 ARIAVSPDAPQLHFVDQYRQOLIARVTSVEVLDKLGQYLQSOYERVALENTRP 1380
 Db 1321 ARIAVSPDAPQLHFVDQYRQOLIARVTSVEVLDKLGQYLQSOYERVALENTRP 1380
 QY 1381 QMRKFLSLSQSWDRCKDGLYQALKEHTPHLIMELWEKSKKLLPLSS 1429
 CC 1381 QMRKFLSLSQSWDRCKDGLYQALKEHTPHLIMELWEKSKKLLPLSS 1429
 Db 1381 QMRKFLSLSQSWDRCKDGLYQALKEHTPHLIMELWEKSKKLLPLSS 1429

RESULT 3

ID ABG78472 standard; Protein; 1429 AA.

XX ABG78472;

DT 15-NOV-2002 (first entry)

XX Leucine-rich repeat (LRR) domain containing protein #2.

XX Human; human leucine-rich repeat small intestine I; HLRRSII; asthma;
 KM proliferative disorder; gastrointestinal disorder; renal disorder;
 KM neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KM immune system; anemia; human immune deficiency virus; HIV; cancer;
 KM blood coagulation disorder; autoimmune disorder; allergic reaction;
 KM inflammatory condition; cardiovascular disorder; ischaemia;
 KM neurological disorder; infectious disease; cytokine production;
 KM expressed sequence tag; EST.

XX Unidentified.

OS MO200261086-A2.

XX 08-AUG-2002.

PF 20-DEC-2001; 2001MO-US49739.

PR 22-DEC-2000; 2000US-257774P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Feder J, Ramanathan C, Mintier G;

DR WPI, 2002-619252/66.

XX New isolated nucleic acid molecules encoding HLRRSII polypeptides, or
 PT their fragments and homologues, useful for preventing, treating and
 PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
 PT or renal disorders

XX Disclosure; Page 229-233; 336pp; English.

XX The invention relates to isolated nucleic acid molecules (I) encoding
 CC human leucine-rich repeat small intestine I (HLRRSII) polypeptides.
 CC The nucleic acid molecules and polypeptides are useful for preventing,
 CC treating and ameliorating medical conditions, such as proliferative,
 CC gastrointestinal, renal, neural, or reproductive disorders, or disorders
 CC related to aberrant calcium regulation or apoptosis modulation, either
 CC directly or indirectly. They are also useful for treating, preventing
 CC and/or diagnosing diseases, disorders and/or conditions of: immune system
 CC by activating or inhibiting the proliferation, differentiation, or

CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
 CC anaemia; immunologic deficiency syndromes; e.g. human immune deficiency
 CC virus (HIV) infection, HTLV-BLV infection; blood coagulation disorders,
 CC e.g. arterial thrombosis; autoimmune disorders; e.g. Addison's disease,
 CC myasthenia gravis; asthma or allergic reactions; inflammatory
 CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
 CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
 CC ischaemia; aneurysms; neurological disorders, e.g. Alzheimer's disease,
 CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
 CC pneumonia, or viral, bacterial, and fungal infections. The HLRRSII
 CC polypeptides are useful for modulating cytokine production, antigen
 CC presentation, or other processes such as boosting immune responses.
 CC ABG78454-ABG78474 represent HLRRSII amino acid sequences and related
 CC amino acid sequences of the invention.

XX Sequence 1429 AA;

Query Match 100.0%; Score 7534; DB 23; Length 1429;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWGRLLACYLEFLKKEKEFOLLANKASRSSSGETPAQPEKTSMEVASYLVNQ 60
 Db 1 MAGGAWGRLLACYLEFLKKEKEFOLLANKASRSSSGETPAQPEKTSMEVASYLVNQ 60
 QY 61 YGEORAMDALHTWEMQGLRSLCAQAOEGAGHSPPFPVSPSEPHLGSPOPTSTAVLMPW 120
 Db 61 YGEORAMDALHTWEMQGLRSLCAQAOEGAGHSPPFPVSPSEPHLGSPOPTSTAVLMPW 120
 QY 121 IHELPAQCTQSGSERVRLRQLPDTSGRMRREISASLLYOLPSSPDHSSQSPNAPST 180
 Db 121 IHELPAQCTQSGSERVRLRQLPDTSGRMRREISASLLYOLPSSPDHSSQSPNAPST 180
 QY 181 AVLGSGWSPQPSLAPEOEAPGTQWPLDETSGIYTYEIRERERESEKGRPPMAVGT 240
 Db 181 AVLGSGWSPQPSLAPEOEAPGTQWPLDETSGIYTYEIRERERESEKGRPPMAVGT 240
 QY 241 PQOAHSLQPHHPWEPVRESLCTWPKNEDFNOKFTOLLILORPHRSQDPLVRSW 300
 Db 241 PQOAHSLQPHHPWEPVRESLCTWPKNEDFNOKFTOLLILORPHRSQDPLVRSW 300
 QY 301 PDVVENRGLHLEIRDLRPGLDTOEPRVILLQGAAGIKSTLARQVKAMRGOLYGR 360
 Db 301 PDVVENRGLHLEIRDLRPGLDTOEPRVILLQGAAGIKSTLARQVKAMRGOLYGR 360
 QY 361 PQHVFYSCRELAQSKVVSIAELIGKDTATPAPIRQLISREBRLFLIDGVDEPGVLQ 420
 Db 361 PQHVFYSCRELAQSKVVSIAELIGKDTATPAPIRQLISREBRLFLIDGVDEPGVLQ 420
 QY 421 EPSSSELCHWSQPOPADALLGSLGKTIIPKASFLITARTALONLIPLEQARWEVIG 480
 Db 421 EPSSSELCHWSQPOPADALLGSLGKTIIPKASFLITARTALONLIPLEQARWEVIG 480
 QY 481 FESSSRKEFYFYPDERQAIRAFRLVKSNEKLMALCIVPWSWLACTLMQMKREKL 540
 Db 481 FESSSRKEFYFYPDERQAIRAFRLVKSNEKLMALCIVPWSWLACTLMQMKREKL 540
 QY 541 TLTSKTTTTLCHYLAQALQAPLGFQDLNLSLAEGIWOKKTLFSPDRLRGHGDGI 600
 Db 541 TLTSKTTTTLCHYLAQALQAPLGFQDLNLSLAEGIWOKKTLFSPDRLRGHGDGI 600
 QY 601 ISTFLMGILQHPHPIPLSYSTFHLCPQEPFAAMSYLDEKGRGKSNCTIDLEKLEAY 660
 Db 601 ISTFLMGILQHPHPIPLSYSTFHLCPQEPFAAMSYLDEKGRGKSNCTIDLEKLEAY 660
 QY 661 GHHGLGASTTFPLGLLDEGEREMENIFHCRLSQGRMLQWVPSQLLQPHSLESJH 720
 Db 661 GHHGLGASTTFPLGLLDEGEREMENIFHCRLSQGRMLQWVPSQLLQPHSLESJH 720
 QY 721 CLVETENKFTLVQVNAHFEMQCVETDNELLVCTFCIKFSRHVKQLQIEGRORSTWS 780
 Db 721 CLVETENKFTLVQVNAHFEMQCVETDNELLVCTFCIKFSRHVKQLQIEGRORSTWS 780

QY 781 PTMVLFRVAVPTDAYWQILFSLVKTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
 DB 781 PTMVLFRVAVPTDAYWQILFSLVKTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
 QY 841 TLRIAGGGLTAEDCKDLAFGLRANQTLTELDLSENVLTDAAGAKLQRLRPSCKLQRLQ 900
 DB 841 TLRIAGGGLTAEDCKDLAFGLRANQTLTELDLSENVLTDAAGAKLQRLRPSCKLQRLQ 900
 QY 901 LVSCGLTSDCCQDLASVLSASPSLKEHLDLQNNLDVGVNLLCGELHNPACKLIRLGLDQ 960
 DB 901 LVSCGLTSDCCQDLASVLSASPSLKEHLDLQNNLDVGVNLLCGELHNPACKLIRLGLDQ 960
 QY 961 TTLSDEMRQELRALQEKPOLIFSRKPSVMTPTGDLTGBMNSNTSLKQRLGSERA 1020
 DB 961 TTLSDEMRQELRALQEKPOLIFSRKPSVMTPTGDLTGBMNSNTSLKQRLGSERA 1020
 QY 1021 ASHVAQNLKLDVSKIFPIAEIAESSPEVVELLCVPSPAQGDILHTKPLGTDDFW 1080
 DB 1021 ASHVAQNLKLDVSKIFPIAEIAESSPEVVELLCVPSPAQGDILHTKPLGTDDFW 1080
 QY 1081 GPTGPVATEVVDKKNLYRHHFPAVGSYRMPNTGLCFVMEKAVTVIEFCWMDQFLGEIN 1140
 DB 1081 GPTGPVATEVVDKKNLYRHHFPAVGSYRMPNTGLCFVMEKAVTVIEFCWMDQFLGEIN 1140
 QY 1141 PSHMVAAGPLDLIKAEPAVEAVHLPFVALQSGHVDTSLFQMAHFKBEGMLKEDPARV 1200
 DB 1141 PSHMVAAGPLDLIKAEPAVEAVHLPFVALQSGHVDTSLFQMAHFKBEGMLKEDPARV 1200
 QY 1201 ELHHIVLENSFSPFLGVLKMHNAALFIPVTSVLLYHRHREPEVTHLYIPSDCSIR 1260
 DB 1201 ELHHIVLENSFSPFLGVLKMHNAALFIPVTSVLLYHRHREPEVTHLYIPSDCSIR 1260
 QY 1261 KEHELCTRSRGEEDLSEFPYVGHGSGIRLOVKKKDETLVWELVYRPGDIMPATTLIP 1320
 DB 1261 KEHELCTRSRGEEDLSEFPYVGHGSGIRLOVKKKDETLVWELVYRPGDIMPATTLIP 1320
 QY 1321 ARIAVPSPLAPOLHFPVDOYREOLIRAVTSVEVVLKLGQVLSQBOYERVLAENTRPS 1380
 DB 1321 ARIAVPSPLAPOLHFPVDOYREOLIRAVTSVEVVLKLGQVLSQBOYERVLAENTRPS 1380
 QY 1381 QMRKLFSLQSWDRCKDGLYQALKEHPHLMELMEKSGKGLPLSS 1429
 DB 1381 QMRKLFSLQSWDRCKDGLYQALKEHPHLMELMEKSGKGLPLSS 1429
 RESULT 4
 ID ABB77916 standard; protein; 1429 AA.
 AC ABB77916;
 DT 07-OCT-2002 (first entry)
 DE Human leucine-rich repeat domain containing protein KIA0926.
 XX Human; leucine-rich repeat; HLRRM1; proliferative disorder;
 KW immune condition; apoptosis; signal transduction; autoimmune disease;
 KW haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
 KW cardiovascular disorder; neurological disease; pheomone;
 KW pulmonary disease; chronic obstructive pulmonary disease;
 KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
 KW haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
 KW inflammatory disorder; systemic lupus erythematosus;
 KW cardiovascular disease; cancer; KIA0926.
 XX Homo sapiens.
 OS
 PN MO200252011-A2.
 PD 04-JUL-2002.
 PF 20-DEC-2001; 2001WO-US49740.

PR 22-DEC-2000; 2000US-257773P.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 XX Feder J, Ramanathan C, Mintier G;
 XX WPL; 2002-566676/60.
 DR
 XX
 PT New HLRRM1 nucleic acids for preventing, treating or ameliorating e.g.
 PT proliferative disorders, immune conditions, a disorder related to
 PT aberrant apoptosis modulation or developmental disorders -
 XX
 PS Example 1; Page 362-366; 371pp; English.
 XX
 CC The present sequence represents a human leucine-rich repeat domain
 CC containing protein, which was used as a probe to search for leucine-rich
 CC repeat containing protein HLRRM1. HLRRM1 polypeptides and
 CC polynucleotides are useful for preventing, treating or ameliorating a
 CC medical condition such as a proliferative disorder, immune condition,
 CC a disorder related to aberrant apoptosis modulation, either directly or
 CC indirectly, and in modulating signal transduction activity in various
 CC cells, tissue and organisms. They are also useful for treating,
 CC preventing, or diagnosing diseases of haematopoietic cells, autoimmune
 CC disease, graft-versus-host disease, allergic conditions (e.g. asthma),
 CC cardiovascular disorders, and neurological diseases, and for increasing
 CC the organisms' ability to synthesize and/or release pheomones. The
 CC polypeptide may also be used in treating, preventing or ameliorating
 CC pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic
 CC rhinitis, or bronchial hyperresponsiveness), reproductive disease,
 CC haematopoietic disease, platelet disorders (e.g. Bernard-Soulier
 CC syndrome), non-infectious disorders (e.g. innate immunity to bacterial
 CC pathogens, or adaptive immune response), immune and inflammatory
 CC disorders (e.g. systemic lupus erythematosus), cardiovascular diseases
 CC and cancers. HLRRM1 nucleic acids may further be used in chromosome
 CC identification or mapping, as a chromosome marker, as molecular weight
 CC markers, as diagnostic probes, in gene therapy, in raising anti-DNA
 CC antibodies, or as antigens for eliciting immune responses.
 XX
 SQ Sequence 1429 AA;
 Query Match 100.0%; Score 7534; DB 23; Length 1429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGGAMGRILACYLFLKKEELKEFOLLANKAHSRSSGGTPPAQPEKTSQMEVASTYVAQ 60
 DB 1 MAGGAMGRILACYLFLKKEELKEFOLLANKAHSRSSGGTPPAQPEKTSQMEVASTYVAQ 60
 QY 61 YGEORAWDLALHTWEOGMLSLCAQAOEGAGHSFPYSPSEPHLAGSPQPTSTAVLMPW 120
 DB 61 YGEORAWDLALHTWEOGMLSLCAQAOEGAGHSFPYSPSEPHLAGSPQPTSTAVLMPW 120
 QY 121 IHELPACTQGSERRVLRQLPDTSGRRREISASLYQALPSSPDHSPQSPENPATST 180
 DB 121 IHELPACTQGSERRVLRQLPDTSGRRREISASLYQALPSSPDHSPQSPENPATST 180
 QY 181 AVTGSWCSPPQSLAPRPGAPGTOMFLDTSGLYTTETIEREREKSEKGRPPAAVGT 240
 DB 181 AVTGSWCSPPQSLAPRPGAPGTOMFLDTSGLYTTETIEREREKSEKGRPPAAVGT 240
 QY 241 PPOAHTSLQHHHPWESVVESTCSMPKNEDEPNQFTOLLQRPHPQSODLVGRSW 300
 DB 241 PPOAHTSLQHHHPWESVVESTCSMPKNEDEPNQFTOLLQRPHPQSODLVGRSW 300
 QY 301 PDVYENRGLIETRDIFGPGDLTOEERIVILQAGAGIGKSTLARQYKAMGRGQLYGR 360
 DB 301 PDVYENRGLIETRDIFGPGDLTOEERIVILQAGAGIGKSTLARQYKAMGRGQLYGR 360
 QY 361 FOHVFYSCRELAKQSKVSLAELIGKDGATTPAPIRQILSRPERLLFTLGVDEPQVQLQ 420
 DB 361 FOHVFYSCRELAKQSKVSLAELIGKDGATTPAPIRQILSRPERLLFTLGVDEPQVQLQ 420
 QY 421 EPSSSELCLHWSQPPQPADALLGSLGKTLTPEASFLTARTALQNLPLSLQARWVEVLG 480

421 EPSSRLCLHNSQOPADALLGSLGKTLTLPASLTIRKTTALONLTPSLQANRNVVLG 480
481 FSSSSRRRYFYFTDERQAIKAEFLVKSNEKLMALCLVWVSWLACTCIMOQKRREKL 540
481 FSSSSRRRYFYFTDERQAIKAEFLVKSNEKLMALCLVWVSWLACTCIMOQKRREKL 540
541 TLNKKTTTTCLNHLAQLAQAQPIGPRQLRCLSLAAGIMQKTLFSPDDLRKGLDGA1 600
541 TLNKKTTTTCLNHLAQLAQAQPIGPRQLRCLSLAAGIMQKTLFSPDDLRKGLDGA1 600
601 ISTEFLKGILOEHIPIPSYSFIHLCPOEFPAASVLEDEKGRKSNCTIIDEKTEAY 660
601 ISTEFLKGILOEHIPIPSYSFIHLCPOEFPAASVLEDEKGRKSNCTIIDEKTEAY 660
661 GINGHFGASTTRFLGLLSDEGEREMENIPIHCRLSQGRNIMQWVPSLQILLQPHSLSLH 720
661 GINGHFGASTTRFLGLLSDEGEREMENIPIHCRLSQGRNIMQWVPSLQILLQPHSLSLH 720
721 CLVETRKKTFLTOVMAHFEEMQMCVETDMLVCTFCIKSRHAKYQILIEGRQHRSTWS 780
721 CLVETRKKTFLTOVMAHFEEMQMCVETDMLVCTFCIKSRHAKYQILIEGRQHRSTWS 780
781 PTWVVLFRWVPTDAYWQILPSVLKATRNKKELDLSGNSLSHSAVKSICKTLRRPRCLLE 840
781 PTWVVLFRWVPTDAYWQILPSVLKATRNKKELDLSGNSLSHSAVKSICKTLRRPRCLLE 840
841 TURLAGGCLTAEDCKDLAFGLRANQTLTELDLSPNVLTDAAGAKLQRLRPSCKLQRLQ 900
841 TURLAGGCLTAEDCKDLAFGLRANQTLTELDLSPNVLTDAAGAKLQRLRPSCKLQRLQ 900
901 LVSGGLTSDCCODLAVLSASPSLKELDLOQNNIDVGVRLCEGLRHPACKLIRLGLDQ 960
901 LVSGGLTSDCCODLAVLSASPSLKELDLOQNNIDVGVRLCEGLRHPACKLIRLGLDQ 960
961 TTLSDEKROELRALEOKRPOLLIFSRKRPVMTPEGLDTGEMSNSTSLKROLGSERA 1020
961 TTLSDEKROELRALEOKRPOLLIFSRKRPVMTPEGLDTGEMSNSTSLKROLGSERA 1020
1021 ASHVAQANLKLIDVSKIFPIAIEASSPEVAVPELLCVBPASQGLHTKPLGTDDEFW 1080
1021 ASHVAQANLKLIDVSKIFPIAIEASSPEVAVPELLCVBPASQGLHTKPLGTDDEFW 1080
1081 GPFGPVATEVVDKRNLYRHFVPAAGSYRWMTGELCFMRAAVVVEIFCWMDFLGEIN 1140
1081 GPFGPVATEVVDKRNLYRHFVPAAGSYRWMTGELCFMRAAVVVEIFCWMDFLGEIN 1140
1141 PSHSMVAGPLLDIKAEPAVEAHLPHFVALQSGHVDTSLFQNAHFKEEGMLLEKPARV 1200
1141 PSHSMVAGPLLDIKAEPAVEAHLPHFVALQSGHVDTSLFQNAHFKEEGMLLEKPARV 1200
1201 ELHHIVLENPSFPLGVLLKMIHNAARFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
1201 ELHHIVLENPSFPLGVLLKMIHNAARFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
1261 KELELCRSPBEDOLFSEFYVGHGSGIRLOVOKKOBTLVWELVYRPGDMLPMTTLIP 1320
1261 KELELCRSPBEDOLFSEFYVGHGSGIRLOVOKKOBTLVWELVYRPGDMLPMTTLIP 1320
1321 ARIAVSPPLDAPOLHFDVQYREOLIAVTSVEVLLDLHGQVLSOEYERVLLENTRPS 1380
1321 ARIAVSPPLDAPOLHFDVQYREOLIAVTSVEVLLDLHGQVLSOEYERVLLENTRPS 1380
1381 QMRKLFSLISQSWDRCKDGLVQALKETHPHILIMELMEKSKKGLPLSS 1429
1381 QMRKLFSLISQSWDRCKDGLVQALKETHPHILIMELMEKSKKGLPLSS 1429

RESULT 5
ABG71631
ID ABG71631 standard; Protein; 1429 AA.
XX
AC ABG71631;

XX 09-JAN-2003 (first entry)
XX Human caspase recruitment domain-7 (CARD-7).
XX Human caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
XX pseudocyt; ICEBERG; cell growth; cell death; inflammation;
XX apoptosis; caspase activation; cancer; follicular lymphoma;
XX leukemia; melanoma; colon cancer; lung carcinoma; viral infection;
XX autoimmune disease; systemic lupus erythematosus; reactive arthritis;
XX human immunodeficiency virus infection; HIV infection; AIDS;
XX Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
XX myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
XX insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
XX graft rejection; allergic rhinitis; food allergy; conjunctivitis;
XX glomerular nephritis; cytostatic; virocid; immunosuppressive;
XX dermatological; nephrotoxic; neuroprotective; cardiant.
OS Homo sapiens.
XX
XX US2002128219-A1.
XX
XX 12-SEP-2002.
XX
XX 15-AUG-2001; 2001US-0931071.
XX
XX 27-OCT-1999; 99US-0428252.
XX
XX (BERT/) BERTIN J.
XX (ALNE/) ALNEMRI E. S.
XX Bertin J, Alnemri ES;
XX WPI; 2003-028968/02.
XX N-PSDB; ABS55497.
XX
XX Assays for identifying compound that modulates the interaction of
XX caspase recruitment domain-8 with a CARD-8 ligand or a compound that
XX modulates activity of CARD-8
XX
XX Disclosure; Fig 1; 49pp; English.
XX
XX The present invention relates to methods of identifying compounds
XX that regulate caspase activity using caspase recruitment domain-7
XX (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
XX method for identifying a compound that modulates the interaction
XX between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
XX converting enzyme (pseudocyt) or ICEBERG is disclosed. CARD-7 and
XX CARD-8 molecules are useful as modulating agents in regulating a
XX variety of cellular processes including cell growth, cell death, and
XX inflammation. The methods of the invention are useful for identifying
XX compounds that have the ability to increase/decrease apoptosis, or
XX comprise the ability to induce caspase activation. The methods are
XX useful for treating a disorder associated with inappropriate apoptosis
XX or inappropriate inflammation. The methods are useful for treating
XX disorders associated with an unduly low rate of apoptosis such
XX as cancer (preferably follicular lymphoma, chronic myelogenous
XX leukemia, melanoma, colon cancer, lung carcinoma, etc), viral
XX infections, autoimmune diseases caused by low levels of apoptosis
XX (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
XX and arthritis). The methods are also useful for treating disorders with
XX unduly high rates of apoptosis such as human immunodeficiency
XX virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
XX amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
XX muscular atrophy, various forms of cerebellar degeneration, anaemia
XX associated with chronic disease, aplastic anaemia, chronic neutropenia,
XX myelodysplastic syndromes, myocardial infarction, stroke, and
XX various inflammatory disorders (e.g. Crohn's disease, reactive
XX arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
XX psoriasis, graft rejection, allergic rhinitis, food allergies,
XX conjunctivitis, glomerular nephritis, etc). The present sequence
XX represents human CARD-7.

XX Sequence 1429 AA;
 Query Match 100.0%; Score 7534; DB 24; Length 1429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAGRLACIYEFLEKKEELKEFOLLANKAHSSSGSTPAQPEKTSQMEVASYVAQ 60
 DB 1 MAGGAGRLACIYEFLEKKEELKEFOLLANKAHSSSGSTPAQPEKTSQMEVASYVAQ 60
 QY 61 YGQRAMDLALHTWQGLRSLCAQAGSGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 120
 DB 61 YGQRAMDLALHTWQGLRSLCAQAGSGSGSPSPSPSPSPSPSPSPSPSPSPSPSPSP 120
 QY 121 IHELPAQCTGSSRRVRLQPLDTSGRMRREISASLLVQALPSSPDHSPSOESPNATST 180
 DB 121 IHELPAQCTGSSRRVRLQPLDTSGRMRREISASLLVQALPSSPDHSPSOESPNATST 180
 QY 181 AVUGSGSPQPSLAPREQAPGTOPLDTSQTYTEIREREREKSEKRPMAAVVGT 240
 DB 181 AVUGSGSPQPSLAPREQAPGTOPLDTSQTYTEIREREREKSEKRPMAAVVGT 240
 QY 241 PPOAHTSILOPHHPMBESVRESICSTWPMKNEPNOKFTQLLLQRPSPQDPLVRSN 300
 DB 241 PPOAHTSILOPHHPMBESVRESICSTWPMKNEPNOKFTQLLLQRPSPQDPLVRSN 300
 QY 301 PDVYENRGLHLEIRDLFGPLDTEPRIVILQAGAGIKSTLARQYKAMGRGQLYGR 360
 DB 301 PDVYENRGLHLEIRDLFGPLDTEPRIVILQAGAGIKSTLARQYKAMGRGQLYGR 360
 QY 361 FOHVFYFSCRELAKQSKVSLAEILGKGTATPAIRQILSRPELFLILGVDEPQVNLQ 420
 DB 361 FOHVFYFSCRELAKQSKVSLAEILGKGTATPAIRQILSRPELFLILGVDEPQVNLQ 420
 QY 421 EPSEELCLHWSOPADALGSLGKTLPEASFLIARFTALQNLPSLEQARWVEVLG 480
 DB 421 EPSEELCLHWSOPADALGSLGKTLPEASFLIARFTALQNLPSLEQARWVEVLG 480
 QY 481 FSSSRKREYFYFTDERQAIRFLVKSNEKELMALCLVFWVSLACTCLMOQKREKL 540
 DB 481 FSSSRKREYFYFTDERQAIRFLVKSNEKELMALCLVFWVSLACTCLMOQKREKL 540
 QY 541 TLVSKTTTTCLHVALDALQAPGPDRLDCSLAAGINQKTLBEPDRLRKGDLGAI 600
 DB 541 TLVSKTTTTCLHVALDALQAPGPDRLDCSLAAGINQKTLBEPDRLRKGDLGAI 600
 QY 601 ISPTLKGIIQEHPIPLSYSPFHLCPOEPFAMSVYLEDEKGRGKSHNCIIDEKTLAY 660
 DB 601 ISPTLKGIIQEHPIPLSYSPFHLCPOEPFAMSVYLEDEKGRGKSHNCIIDEKTLAY 660
 QY 661 GIHGLFGASTTRFLGLSLDEGEREMENIFACRLSQRNLMQWPVSLQLLQPHSLSLH 720
 DB 661 GIHGLFGASTTRFLGLSLDEGEREMENIFACRLSQRNLMQWPVSLQLLQPHSLSLH 720
 QY 721 CLYETRNKTFITQVMAHFEKMCVEETDMELVCTFCIKSRHYKQQLIEGRHRSTWS 780
 DB 721 CLYETRNKTFITQVMAHFEKMCVEETDMELVCTFCIKSRHYKQQLIEGRHRSTWS 780
 QY 781 PTWVVLFRWGVVDAYVQILPSVLKVTNRNKELDLSGNSLSHSAVKSILCTLRPRCLLE 840
 DB 781 PTWVVLFRWGVVDAYVQILPSVLKVTNRNKELDLSGNSLSHSAVKSILCTLRPRCLLE 840
 QY 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKILCQRLRQPSCKLRLQ 900
 DB 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKILCQRLRQPSCKLRLQ 900
 QY 901 LVSGGLTSDCCODLAVSLASPSLKEILDQONNLDVGVRLLCGLHHPACKLIRLGLDQ 960
 DB 901 LVSGGLTSDCCODLAVSLASPSLKEILDQONNLDVGVRLLCGLHHPACKLIRLGLDQ 960
 QY 961 TTLSDEMRQELRALEOKRPOLLISRRKPSVMTPTGELDTGEMNSTSLKRORLGSERA 1020
 DB 961 TTLSDEMRQELRALEOKRPOLLISRRKPSVMTPTGELDTGEMNSTSLKRORLGSERA 1020

DB 961 TTLSDEMRQELRALEOKRPOLLISRRKPSVMTPTGELDTGEMNSTSLKRORLGSERA 1020
 QY 1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEVVEVVELLCVPSPASQGDLTHTPLGTTDDFW 1080
 DB 1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEVVEVVELLCVPSPASQGDLTHTPLGTTDDFW 1080
 QY 1081 GPTGPVATEVVDKRLVYHFPVAGSYRMPNGLCFWMEAVTVEIEFCVMDQFLGEIN 1140
 DB 1081 GPTGPVATEVVDKRLVYHFPVAGSYRMPNGLCFWMEAVTVEIEFCVMDQFLGEIN 1140
 QY 1141 PHSMMVAGPLDITKAEPGAVEAVHLPHFVALQGHVDTSLFQMAHPEKEGMLEKPARV 1200
 DB 1141 PHSMMVAGPLDITKAEPGAVEAVHLPHFVALQGHVDTSLFQMAHPEKEGMLEKPARV 1200
 QY 1201 ELHRIVLNPSFSPGLVLLKMINALRPIVTSVLLYHRVHBEVTFHLIIPSDCSIR 1260
 DB 1201 ELHRIVLNPSFSPGLVLLKMINALRPIVTSVLLYHRVHBEVTFHLIIPSDCSIR 1260
 QY 1261 KELELCYRSPGEOQLSEFVGHLSGIRLQVQDKDETLWMBALVYKPGMLPATTLIP 1320
 DB 1261 KELELCYRSPGEOQLSEFVGHLSGIRLQVQDKDETLWMBALVYKPGMLPATTLIP 1320
 QY 1321 ARIAVSPDLAPQLLHFVDQYREQLIARVTSVEVLDKLHGVLQSGQYERVLAENTRPS 1380
 DB 1321 ARIAVSPDLAPQLLHFVDQYREQLIARVTSVEVLDKLHGVLQSGQYERVLAENTRPS 1380
 QY 1381 QMRKLFSLSGSMRCKCKDGLYQALKETHPHLIMELWEKSKGLPLPSS 1429
 DB 1381 QMRKLFSLSGSMRCKCKDGLYQALKETHPHLIMELWEKSKGLPLPSS 1429

RESULT 6
 ABG71633 standard; Protein; 1429 AA.
 ID ABG71633
 XX
 AC ABG71633;
 XX
 DT 10-JAN-2003 (first entry)
 XX
 DE Human caspase recruitment domain-7 (CARD-7).
 XX
 KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudotice; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukemia; melanoma; colon cancer; lung cancer; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cycostatic; viroicide; immunosuppressive;
 KW dermatological; nephrotoxic; neuroprotective; cardiac.
 XX
 OS Homo sapiens.
 XX
 PN US2002128198-A1.
 XX
 PD 12-SEP-2002.
 XX
 PF 27-NOV-2001; 2001US-0996617.
 XX
 PR 28-JUN-1999; 99US-0340620.
 PR 27-OCT-1999; 99US-0428252.
 PR 15-AUG-2001; 2001US-0931071.
 XX
 PA (BERT/) BERTIN J.
 XX
 PI Bertin J;
 XX
 DR WPI; 2003-028967/02.
 DR N-PSDB; ABS56030.

XX Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT CARD-7 and CARD-5 in presence of test compound, measuring their
PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PS altered -
XX Disclosure; Fig 1; 43pp; English.
XX The present invention relates to methods of identifying compounds
CC that regulate caspase activity using caspase recruitment domain-7
CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC method for identifying a compound that modulates the interaction
CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC converting enzyme (pseudoc1) or ICEBERG is disclosed. CARD-7 and
CC CARD-8 molecules are useful as modulating agents in regulating a
CC variety of cellular processes including cell growth, cell death, and
CC inflammation. The methods of the invention are useful for identifying
CC compounds that have the ability to increase/decrease apoptosis, or
CC comprise the ability to induce caspase activation. The methods are
CC useful for treating a disorder associated with inappropriate apoptosis
CC or inappropriate inflammation. The methods are useful for treating
CC disorders associated with an undetectable low rate of apoptosis such
CC as cancer (preferably follicular lymphoma, chronic myelogenous
CC leukaemia, melanoma, colon cancer, lung carcinoma, etc.) viral
CC infections, autoimmune diseases caused by low levels of apoptosis
CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC and arthritis). The methods are also useful for treating disorders with
CC undetectably high rates of apoptosis such as human immunodeficiency
CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC muscular atrophy, various forms of cerebellar degeneration, anaemia,
CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
CC myelodysplastic syndromes, myocardial infarction, stroke, and
CC various inflammatory disorders (e.g. Crohn's disease, reactive
CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC psoriasis, graft rejection, allergic rhinitis, food allergies,
CC conjunctivitis, glomerular nephritis, etc). The present sequence
CC represents human CARD-7.
XX
XX Sequence 1429 AA:
SQ
Query Match 100.0%; Score 7534; DB 24; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSSSSGGTTPAOKETSGMEVASYLVAQ 60
Db 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSSSSGGTTPAOKETSGMEVASYLVAQ 60
QY 61 YGEORANDLALHTWEQWGLSLCAQAOEGAGHSPSPYSPBPHLGSPOPTSTAVLMPW 120
Db 61 YGEORANDLALHTWEQWGLSLCAQAOEGAGHSPSPYSPBPHLGSPOPTSTAVLMPW 120
QY 121 IHEIPACCTGSSERRVRLPDTSGRRWREISASLALYQALPSSPDHSSPOESNATPST 180
Db 121 IHEIPACCTGSSERRVRLPDTSGRRWREISASLALYQALPSSPDHSSPOESNATPST 180
QY 181 AVLSGWSPPQPSLAPREOAPGTQWPLDTSIGIYTERERERERERERERERERERER 240
Db 181 AVLSGWSPPQPSLAPREOAPGTQWPLDTSIGIYTERERERERERERERERERERER 240
QY 241 PPOAHTSLQPHHHPWBSVRESLCTWPKNEDFNQFTQLLIQPHPSODPLVGRSW 300
Db 241 PPOAHTSLQPHHHPWBSVRESLCTWPKNEDFNQFTQLLIQPHPSODPLVGRSW 300
QY 301 PDVYENRGLHIEIRDFGPGLDTOBPRIVILQGAAGIGSTLARQYKAMGRGQLGDR 360
Db 301 PDVYENRGLHIEIRDFGPGLDTOBPRIVILQGAAGIGSTLARQYKAMGRGQLGDR 360
QY 361 FOHVYFSCRELASQKSVSLAELIGKDGATPAPRIOLISPERLLFILDGVDPERGVLQ 420
Db 361 FOHVYFSCRELASQKSVSLAELIGKDGATPAPRIOLISPERLLFILDGVDPERGVLQ 420

QY 421 EBSSELCHMSOPQPADLIGSLGKTLIPBASFLITARTALONLIPSEORWBEVLG 480
Db 421 EBSSELCHMSOPQPADLIGSLGKTLIPBASFLITARTALONLIPSEORWBEVLG 480
QY 481 FSSSRKREYRYFTBEROARAFRLVKSNEKMAELCLVWWSMACTCMQMKREKL 540
Db 481 FSSSRKREYRYFTBEROARAFRLVKSNEKMAELCLVWWSMACTCMQMKREKL 540
QY 541 TLTSKTTTLCLHYLAQALQAPRLCSLAAGIWOXKTLFSPDDLRRKGLDGA1 600
Db 541 TLTSKTTTLCLHYLAQALQAPRLCSLAAGIWOXKTLFSPDDLRRKGLDGA1 600
QY 601 ISFPLKMGILQEPRIPLSYFHLQCFEPFAMSVYLEDKGGKSNCTIDLEKLEAY 660
Db 601 ISFPLKMGILQEPRIPLSYFHLQCFEPFAMSVYLEDKGGKSNCTIDLEKLEAY 660
QY 661 GINGLFCASRTTRFLGLSDEGEREMENPHCLSGORNMOMVPSQLLOPHSLSLH 720
Db 661 GINGLFCASRTTRFLGLSDEGEREMENPHCLSGORNMOMVPSQLLOPHSLSLH 720
QY 721 CLVETRNKFTLQVMAHFEEMGMCVETDMLVCTFCIKFSRHVKQLLIEGRQHSWTS 780
Db 721 CLVETRNKFTLQVMAHFEEMGMCVETDMLVCTFCIKFSRHVKQLLIEGRQHSWTS 780
QY 781 PMWVLPFRWVPTDAYWQILFSLVKTTRNKELDLSGNSLSHSAVSLCTTARPRCLLE 840
Db 781 PMWVLPFRWVPTDAYWQILFSLVKTTRNKELDLSGNSLSHSAVSLCTTARPRCLLE 840
QY 841 TLRLAGGGLAEBCKDLAPFLRANOTLTLELDLSPNVLTDPAGNLCORLROPSCKLORLQ 900
Db 841 TLRLAGGGLAEBCKDLAPFLRANOTLTLELDLSPNVLTDPAGNLCORLROPSCKLORLQ 900
QY 901 LVSCGLTSDCCDQDASVLSASPSLKELDLQONNLDVGVNLLCEGLRHAPCKLIRGLDQ 960
Db 901 LVSCGLTSDCCDQDASVLSASPSLKELDLQONNLDVGVNLLCEGLRHAPCKLIRGLDQ 960
QY 961 TTLSDEKROELALBOEKPOLLFSSRRKPSVMTPTGLDTGEMSNSSLSKROLSERA 1020
Db 961 TTLSDEKROELALBOEKPOLLFSSRRKPSVMTPTGLDTGEMSNSSLSKROLSERA 1020
QY 1021 ASHVAQNLKLDLVSXIFPAIAIESSPEVNEVELCVSPSPASQGDILTKPLGTDDDFW 1080
Db 1021 ASHVAQNLKLDLVSXIFPAIAIESSPEVNEVELCVSPSPASQGDILTKPLGTDDDFW 1080
QY 1081 GPTGPVATEVVDKEXLRYVHPVAGSYRWPNTGLCFVMEAVTVIEFCVMDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKEXLRYVHPVAGSYRWPNTGLCFVMEAVTVIEFCVMDQFLGEIN 1140
QY 1141 PHSMMVAGPLDILKABPGVAEAVNHPHVALOGGHVDTSLFQMAHKEEGMLLEKPARV 1200
Db 1141 PHSMMVAGPLDILKABPGVAEAVNHPHVALOGGHVDTSLFQMAHKEEGMLLEKPARV 1200
QY 1201 ELHRIYLENPSFPLGLVLMKINNALRFTIPVTSVLLYHRVHEEVTFHLYLIPDSISR 1260
Db 1201 ELHRIYLENPSFPLGLVLMKINNALRFTIPVTSVLLYHRVHEEVTFHLYLIPDSISR 1260
QY 1261 KELBELCYRSPGSDQFLSEFVYVGLSGGIRLVOKKXDETLWEALYKPGDLPATLIP 1320
Db 1261 KELBELCYRSPGSDQFLSEFVYVGLSGGIRLVOKKXDETLWEALYKPGDLPATLIP 1320
QY 1321 ARIAVSPDLAPOLLHFDVQRYRQOLARVTSVAVLADKLHGQVLSQGYRVLAEENRPS 1380
Db 1321 ARIAVSPDLAPOLLHFDVQRYRQOLARVTSVAVLADKLHGQVLSQGYRVLAEENRPS 1380
QY 1381 QMRKLSLSQSWDKCKDGLYQALKETHPHLIMELWEGSKKGLPLSS 1429
Db 1381 QMRKLSLSQSWDKCKDGLYQALKETHPHLIMELWEGSKKGLPLSS 1429
RESULT 7
AA72669
ID AA72669 standard; Protein; 1473 AA.

AC AAY72669;
 XX
 DT 31-MAY-2001 (first entry)
 DE Human NB-ARC and CARD containing protein (NAC) beta isoform.
 XX
 XX Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 XX caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KM cysteine aspartyl protease; apoptosis; cytokine production;
 KM cytokine receptor; signalling; therapy; inflammatory disorder; sepsis;
 XX fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Location/Qualifiers
 FT 329..547
 FT /label= NB domain
 FT /note= "Nucleotide binding domain, also designated as
 FT NB-ARC domain"
 FT 335..341
 FT /label= Walker_A
 FT /note= "Also designated as P-loop"
 FT 406..414
 FT /label= Walker_B
 FT 809..833
 FT /label= Leucine_rich_repeat_region
 FT 838..862
 FT /label= Leucine_rich_repeat_region
 FT 865..890
 FT /label= Leucine_rich_repeat_region
 FT 895..919
 FT /label= Leucine_rich_repeat_region
 FT 923..947
 FT /label= Leucine_rich_repeat_region
 FT 957..987
 FT /note= "This 31 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) and NAC delta isoform
 FT (AAY72671) due to alternative mRNA splicing"
 FT 1261..1306
 FT /note= "This 45 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) due to alternative
 FT mRNA splicing"
 FT 1079..1364
 FT /note= "TIM-Barrel-like domain"
 FT 1128..1473
 FT /label= CARD-L
 FT /note= "Caspase-associated recruitment domain"
 FT 1128..1261
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1306..1473
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1373..1473
 FT /label= CARD
 FT /note= "Caspase-associated recruitment domain"
 XX
 PN WO200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000MO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 XX (BURN-) BURHAM INST.
 XX
 PA
 PI Reed JC;
 XX
 DR WPI, 2001-183258/18.
 DR N-PSDB; AAD02760.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment

PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 PS Claim 15, Page 133-137; 184pp; English.
 XX
 XX The present sequence is a human NB-ARC and CARD containing protein
 CC (NAC) beta isoform. NAC beta isoform represents the NAC splice
 CC variant in which both the splice regions are present in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukemias.
 CC Note: This sequence is stated as being the same as that shown as
 CC SEQ ID NO:2 (AAY72711) in figure 1A of the specification. However the
 CC sequences differ at several positions.
 XX
 XX Sequence 1473 AA;
 SQ
 Query Match 99.6%; Score 7502; DB 22; Length 1473;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAGGAWRIACYLEFLKKEELKEFOLLANKAHSSSGGTTPAQPEKSGMEVASYLVAQ 60
 DB 1 MAGGAWRIACYLEFLKKEELKEFOLLANKAHSSSGGTTPAQPEKSGMEVASYLVAQ 60
 QY 61 YGEQRAMDALATHWEQNGRLSLCAQOEGAGHSPSPYSSSEHLSSPSTAYLMPW 120
 DB 61 YGEQRAMDALATHWEQNGRLSLCAQOEGAGHSPSPYSSSEHLSSPSTAYLMPW 120
 QY 121 IHELPAGCTQGSRRYLRLPDTSGRWREISASLYOALPSPDHSPQESPNAPTST 180
 DB 121 IHELPAGCTQGSRRYLRLPDTSGRWREISASLYOALPSPDHSPQESPNAPTST 180
 QY 121 IHELPAGCTQGSRRYLRLPDTSGRWREISASLYOALPSPDHSPQESPNAPTST 180
 DB 121 IHELPAGCTQGSRRYLRLPDTSGRWREISASLYOALPSPDHSPQESPNAPTST 180
 QY 181 AVLGWSGSPQPSLAPREQAPCTQWPLDTSGIYVTEIRERRESEKRPMAAVGT 240
 DB 181 AVLGWSGSPQPSLAPREQAPCTQWPLDTSGIYVTEIRERRESEKRPMAAVGT 240
 QY 181 AVLGWSGSPQPSLAPREQAPCTQWPLDTSGIYVTEIRERRESEKRPMAAVGT 240
 DB 181 AVLGWSGSPQPSLAPREQAPCTQWPLDTSGIYVTEIRERRESEKRPMAAVGT 240
 QY 241 PPOAHTSLQPHHHPWSPVRESICSTWPKNEPNOKFTOLLLOPHRPSOPLYKRSW 300
 DB 241 PPOAHTSLQPHHHPWSPVRESICSTWPKNEPNOKFTOLLLOPHRPSOPLYKRSW 300
 QY 241 PPOAHTSLQPHHHPWSPVRESICSTWPKNEPNOKFTOLLLOPHRPSOPLYKRSW 300
 DB 241 PPOAHTSLQPHHHPWSPVRESICSTWPKNEPNOKFTOLLLOPHRPSOPLYKRSW 300
 QY 301 PDVVEENRGHLIRLDFGGLDTQBPRIYILGAAIGIKSTLAROYKEMGRGQLYGR 360
 DB 301 PDVVEENRGHLIRLDFGGLDTQBPRIYILGAAIGIKSTLAROYKEMGRGQLYGR 360
 QY 301 PDVVEENRGHLIRLDFGGLDTQBPRIYILGAAIGIKSTLAROYKEMGRGQLYGR 360
 DB 301 PDVVEENRGHLIRLDFGGLDTQBPRIYILGAAIGIKSTLAROYKEMGRGQLYGR 360
 QY 361 FOHVFFYSCRELAQSKVSLAEILGKGATPAPRIQLSRPRLFLIDGVEPGVLQ 420
 DB 361 FOHVFFYSCRELAQSKVSLAEILGKGATPAPRIQLSRPRLFLIDGVEPGVLQ 420
 QY 361 FOHVFFYSCRELAQSKVSLAEILGKGATPAPRIQLSRPRLFLIDGVEPGVLQ 420
 DB 361 FOHVFFYSCRELAQSKVSLAEILGKGATPAPRIQLSRPRLFLIDGVEPGVLQ 420
 QY 421 EPSSSELCIHWSPQPPDALIGSLGKTIIPBASFLITARTALONLIPSEQARWEVLG 480
 DB 421 EPSSSELCIHWSPQPPDALIGSLGKTIIPBASFLITARTALONLIPSEQARWEVLG 480
 QY 421 EPSSSELCIHWSPQPPDALIGSLGKTIIPBASFLITARTALONLIPSEQARWEVLG 480
 DB 421 EPSSSELCIHWSPQPPDALIGSLGKTIIPBASFLITARTALONLIPSEQARWEVLG 480
 QY 481 FSESRSRKYRYRYTEROQIRAFRLVKSKEMLALCLVWVSWMLACTCLMOQMKREKL 540
 DB 481 FSESRSRKYRYRYTEROQIRAFRLVKSKEMLALCLVWVSWMLACTCLMOQMKREKL 540
 QY 481 FSESRSRKYRYRYTEROQIRAFRLVKSKEMLALCLVWVSWMLACTCLMOQMKREKL 540
 DB 481 FSESRSRKYRYRYTEROQIRAFRLVKSKEMLALCLVWVSWMLACTCLMOQMKREKL 540
 QY 541 TLTSKTTTTLCLHYLAQALQAPLQPLRDLCSLAAGIWKTKLSPDDLRLRGHLDGAI 600
 DB 541 TLTSKTTTTLCLHYLAQALQAPLQPLRDLCSLAAGIWKTKLSPDDLRLRGHLDGAI 600
 QY 541 TLTSKTTTTLCLHYLAQALQAPLQPLRDLCSLAAGIWKTKLSPDDLRLRGHLDGAI 600
 DB 541 TLTSKTTTTLCLHYLAQALQAPLQPLRDLCSLAAGIWKTKLSPDDLRLRGHLDGAI 600
 QY 601 ISTFLKGIIOEHPILPLSYFHLCTOEPPAANSYLLDEKGGKSNCTIIDLEKTEAY 660
 DB 601 ISTFLKGIIOEHPILPLSYFHLCTOEPPAANSYLLDEKGGKSNCTIIDLEKTEAY 660
 QY 601 ISTFLKGIIOEHPILPLSYFHLCTOEPPAANSYLLDEKGGKSNCTIIDLEKTEAY 660
 DB 601 ISTFLKGIIOEHPILPLSYFHLCTOEPPAANSYLLDEKGGKSNCTIIDLEKTEAY 660
 QY 661 GHHGLFGASTTRRLGLSLDEGEREMENFHCRLSGGRNLQWVPSLIQLLPHSLSLH 720

```

Db 661 GINGLGASTRTRLGLISBGEREMENIFCHGLSCGRNLMQVPSQLLIQPHSLH 720
Qy 721 CLVETNKTFLLIQVMAHFEEMGMCVETDMELLVCTFCIKFSRVKQLQLIEGRHSTWS 780
Db 721 CLVETNKTFLLIQVMAHFEEMGMCVETDMELLVCTFCIKFSRVKQLQLIEGRHSTWS 780
Qy 781 PTWVLFRRWVPTDAMQQLPSVLKATRNKEIDLSGNSLSHSAVSLICTLARPPCLIE 840
Db 781 PTWVLFRRWVPTDAMQQLPSVLKATRNKEIDLSGNSLSHSAVSLICTLARPPCLIE 840
Qy 841 TRLACGSLTAEDCKDLAFGLRANQTLTELDLSFNVLTDGAHLCORLRQPSCKQLRLQ 900
Db 841 TRLACGSLTAEDCKDLAFGLRANQTLTELDLSFNVLTDGAHLCORLRQPSCKQLRLQ 900
Qy 901 LVSCGLTSDCCODLAVSLASPSIKELIDLQNNLDVGVRLCEGLRHPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLAVSLASPSIKELIDLQNNLDVGVRLCEGLRHPACKLIRGLDQ 960
Qy 961 TTLSDEMROELRLBOEKPOLLI PSRRKPSVMTPTGLDGMGMSNSTSLKQRLSERA 1020
Db 961 TTLSDEMROELRLBOEKPOLLI PSRRKPSVMTPTGLDGMGMSNSTSLKQRLSERA 1020
Qy 1021 ASHVAQANLKLIDVSKIFPIAEIAESSPEVVEVELLCVPSPASQGLHTKPLGTDDFW 1080
Db 1021 ASHVAQANLKLIDVSKIFPIAEIAESSPEVVEVELLCVPSPASQGLHTKPLGTDDFW 1080
Qy 1081 GPTGPVATEVVDKEXKILYRVHFPVAGSYRMPNTGLCFVMEAVTVEIIFCVMDOFLGEIN 1140
Db 1081 GPTGPVATEVVDKEXKILYRVHFPVAGSYRMPNTGLCFVMEAVTVEIIFCVMDOFLGEIN 1140
Qy 1141 PHSMMVAGPLDILIKAPGAVEAVHLPHFVALQGHVDSLFQMAHKEKGMLLEKPARV 1200
Db 1141 PHSMMVAGPLDILIKAPGAVEAVHLPHFVALQGHVDSLFQMAHKEKGMLLEKPARV 1200
Qy 1201 ELHIIYLENPSFSPGLGVLLKMINALRFIPVTSVLLYHHPBEVTFPHLYLIPDSCIR 1260
Db 1201 ELHIIYLENPSFSPGLGVLLKMINALRFIPVTSVLLYHHPBEVTFPHLYLIPDSCIR 1260
Qy 1261 -----KELELCYRSPGEDOLF 1276
Db 1261 KAIDLEMKFOVRIHKKPPLTPLYMGCRYTSGSGSMLEILPKLELELCYRSPGEDOLF 1320
Qy 1277 SEFYVGHLSGSLILOYKDKDELVWEALVKPGDIMPATLLIPARIAVSPPLDAPQLH 1336
Db 1321 SEFYVGHLSGSLILOYKDKDELVWEALVKPGDIMPATLLIPARIAVSPPLDAPQLH 1380
Qy 1337 FVDQYREQLIARYTSVEVVDKLGCVLSQOYERYVLAENTRPSQMKLFSLSQSMDKRC 1396
Db 1381 FVDQYREQLIARYTSVEVVDKLGCVLSQOYERYVLAENTRPSQMKLFSLSQSMDKRC 1440
Qy 1397 KQGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
Db 1441 KQGLYQALKETHPHLIMELWEKSKKGLPLSS 1473

RESULT 8
AA72711
ID AA72711 standard; Protein; 1473 AA.
XX
AC AA72711;
XX
DT 31-MAY-2001 (first entry)
XX
DE Human NAC beta isoform, alternative version.
XX
KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
cysteine asparaginyl protease; apoptosis; cytokine production;
cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
fibrosis; arthritis; cancer; adenocarcinoma; leukemia.
XX
OS Homo sapiens.

```

```

XX Key Location/Qualifiers
FH Domain 329..547
FT /label= NB domain
FT /note= "Nucleotide binding domain, also designated as
FT NB ARC domain"
FT 329..341
FT /label= Walker A
FT /note= "Also designated as P-loop"
FT 406..414
FT /label= Walker B
FT 809..833
FT /label= Leucine_rich_repeat_region
FT 838..862
FT /label= Leucine_rich_repeat_region
FT 865..890
FT /label= Leucine_rich_repeat_region
FT 895..919
FT /label= Leucine_rich_repeat_region
FT 923..947
FT /label= Leucine_rich_repeat_region
FT 957..987
FT /label= Leucine_rich_repeat_region
FT /note= "This 31 amino acid segment is not found in
FT NAC gamma isoform (AA72670) and NAC delta isoform
FT (AA72671) due to alternative mRNA splicing"
FT 1261..1306
FT /note= "This 45 amino acid segment is not found in
FT NAC gamma isoform (AA72670) due to alternative
FT mRNA splicing"
FT 1079..1364
FT /note= "TIM-Barrel-like domain"
FT 1128..1473
FT /label= CARD-L
FT /note= "Caspase-associated recruitment domain"
FT 1128..1261
FT /label= CARD-S
FT /note= "Caspase-associated recruitment domain"
FT /note= "Caspase-associated recruitment domain"
FT 1298..1305
FT /note= "Encoded by GGGATGCTGGAATACTCCCAAG"
FT 1306..1473
FT /label= CARD-S
FT /note= "Caspase-associated recruitment domain"
FT 1373..1473
FT /label= CARD
FT /note= "Caspase-associated recruitment domain"
PN WO200116170-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000MO-US24152.
XX
PR 01-SEP-1999; 99US-0388221.
PA (BURN-) BURHAM INST.
XX
PI Reed JC;
XX
PI WPI; 2001-183258/18.
XX
DR N-PSDB; AAD02760.
XX
PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT domains, used to produce polypeptides for screening for modulators of
PT apoptosis -
XX
PS Claim 15; Fig 1A; 184bp; English.
XX
CC The present sequence is a human NB-ARC and CARD containing protein
CC (NAC) beta isoform, alternative version. NAC beta isoform represents the
CC NAC splice variant in which both the splice regions are present in the
CC translated polypeptide. NAC protein comprises a nucleotide binding (NB)
CC domain (also referred as NB-ARC domain), a caspase-associated recruitment
CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine

```

CC aspartyl-proteases, are principal effectors of apoptosis, CARD containing
CC NAC proteins are used for screening modulators that modulates apoptosis,
CC cytokine production, cytokine receptor signaling and other cellular
CC processes. NAC can act as an immunogen for the production of polyclonal
CC and monoclonal antibodies. It can also be used to diagnose and treat
CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC pathologies such as adenocarcinomas and leukemias.
CC Note: This sequence is stated as being the same as that shown as
CC SEQ ID NO:2 (See PAM72665) in page 133-137 of the specification. However
CC the sequences differ at several positions.

SQ Sequence 1473 AA;

Query Match	99.6%;	Score 7502;	DB 22;	Length 1473;
-------------	--------	-------------	--------	--------------

Best Local Similarity 97.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY	1	MAGSAGRSLACTLEF,KEBELKEPOLLANKNHSSSSGETPADEKTSQMEVASYVAQ	60
QY	1	MAGSAGRSLACTLEF,KEBELKEPOLLANKNHSSSSGETPADEKTSQMEVASYVAQ	60
QY	1	MAGSAGRSLACTLEF,KEBELKEPOLLANKNHSSSSGETPADEKTSQMEVASYVAQ	60
QY	61	YGEORAMDLALTWMEQMLRSLCAOAOAGHSPPYSPSEBPH,GSPOQTSYAVLMPV	120
QY	61	YGEORAMDLALTWMEQMLRSLCAOAOAGHSPPYSPSEBPH,GSPOQTSYAVLMPV	120
QY	121	IHELPRAGCTQSGERRYLROLPTSGRRMRBISASLLYALBSSSDHSPQOESNAPTST	180
QY	121	IHELPRAGCTQSGERRYLROLPTSGRRMRBISASLLYALBSSSDHSPQOESNAPTST	180
QY	181	AVLGSNGSPQPSLAPREOAPGTQMP,DEPISGITYTTERREBEKEKAPPAVAVGT	240
QY	181	AVLGSNGSPQPSLAPREOAPGTQMP,DEPISGITYTTERREBEKEKAPPAVAVGT	240
QY	241	PROAHTSIOPHHHPWEPVSRESLCTSPWKNEDFNOKETQULLLQRPHPRODPLVKRW	300
QY	241	PROAHTSIOPHHHPWEPVSRESLCTSPWKNEDFNOKETQULLLQRPHPRODPLVKRW	300
QY	301	PRVVEBNRCH,LEIBDLRGGLDQOEBP,IVILQAAIGKSTARQYKEMAGGOLYGD	360
QY	301	PRVVEBNRCH,LEIBDLRGGLDQOEBP,IVILQAAIGKSTARQYKEMAGGOLYGD	360
QY	361	FOHVFFS,CRELQASQKVS,LAELI,IKDQATAPAP,RIQILSRPERIL,FLDGVBPQVQLQ	420
QY	361	FOHVFFS,CRELQASQKVS,LAELI,IKDQATAPAP,RIQILSRPERIL,FLDGVBPQVQLQ	420
QY	421	EPBSECLHWSQOPQADAL,GSLLGKTLT,PEASPLTARTTALONL,PSLEQARWEVLG	480
QY	421	EPBSECLHWSQOPQADAL,GSLLGKTLT,PEASPLTARTTALONL,PSLEQARWEVLG	480
QY	481	FBSSSKEXFYFYF,DERQALTRAP,LYVSNKEWAL,CLVPVNSLACTCLMOQMKREKL	540
QY	481	FBSSSKEXFYFYF,DERQALTRAP,LYVSNKEWAL,CLVPVNSLACTCLMOQMKREKL	540
QY	541	TTTSKTTTTL,CHUYLAOLQALQ,PLPQ,RLDCSL,LAEGIWOKT,LFSPDDL,RLKHGLGAI	600
QY	541	TTTSKTTTTL,CHUYLAOLQALQ,PLPQ,RLDCSL,LAEGIWOKT,LFSPDDL,RLKHGLGAI	600
QY	601	ISTFLKMGILQOHP,RLP,LSYF,HL,CFQEPFAMS,VL,EDGK,SHNC,II,DEKTL,EA	660
QY	601	ISTFLKMGILQOHP,RLP,LSYF,HL,CFQEPFAMS,VL,EDGK,SHNC,II,DEKTL,EA	660
QY	661	GHHG,FGASTP,FL,GL,SLDGBER,ME,NI,PH,CR,LSQGN,LMQW,PS,QL,LLQ,PH,SL,ESLH	720
QY	661	GHHG,FGASTP,FL,GL,SLDGBER,ME,NI,PH,CR,LSQGN,LMQW,PS,QL,LLQ,PH,SL,ESLH	720
QY	721	CLYETRNK,FL,TV,NAH,FEEMQ,MCV,ET,ME,VL,CT,FC,IK,FR,AYK,QL,IEGR,HS,TVS	780
QY	721	CLYETRNK,FL,TV,NAH,FEEMQ,MCV,ET,ME,VL,CT,FC,IK,FR,AYK,QL,IEGR,HS,TVS	780
QY	781	PRMVVL,FRVAV,PTDA,WO,IL,BS,VL,KVTN,KL,BL,SG,NS,LS,HS,VN,SL,CT,LR,PR,CL,LE	840
QY	781	PRMVVL,FRVAV,PTDA,WO,IL,BS,VL,KVTN,KL,BL,SG,NS,LS,HS,VN,SL,CT,LR,PR,CL,LE	840

QY	841	TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDGAGHLQRLQFQPSCKLQRLQ	900
Db	841	TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDGAGHLQRLQFQPSCKLQRLQ	900
QY	901	LVSQGLSPDCODPLASVLSASPSLKELDLQNNLDVGVYLLDEGLRHPACKLIRGLGDC	960
Db	901	LVSQGLSPDCODPLASVLSASPSLKELDLQNNLDVGVYLLDEGLRHPACKLIRGLGDC	960
QY	961	TTLSDEMRQELRALQEKPOLLIFSRRKPSVMTPTBGLDTGEMSNSTSSLKQRLGSERA	1020
Db	961	TTLSDEMRQELRALQEKPOLLIFSRRKPSVMTPTBGLDTGEMSNSTSSLKQRLGSERA	1020
QY	1021	ASHVAQANLKLIDVSKIFPIAETLAEBSSPEVVPEYELLCPVSPASQGDILHTKPIGTDDFW	1080
Db	1021	ASHVAQANLKLIDVSKIFPIAETLAEBSSPEVVPEYELLCPVSPASQGDILHTKPIGTDDFW	1080
QY	1081	GPSPVATVEVDDEKQNLVYRHFPVAGSYRPMNGLCPVMEEAATVTEIEFCVMPQFGEIN	1140
Db	1081	GPSPVATVEVDDEKQNLVYRHFPVAGSYRPMNGLCPVMEEAATVTEIEFCVMPQFGEIN	1140
QY	1141	POHSMWVAGPLDLIKAPGAVEAVHLPHFVALQSGHYDTSLFQMAHFKKEGMILLEKPARV	1200
Db	1141	POHSMWVAGPLDLIKAPGAVEAVHLPHFVALQSGHYDTSLFQMAHFKKEGMILLEKPARV	1200
QY	1201	ELAHIVLENPSPSPGLVLLKMLHNALEFIPVTSVILLYHRVPEEYTFHLYLIPDQCSIR	1260
Db	1201	ELAHIVLENPSPSPGLVLLKMLHNALEFIPVTSVILLYHRVPEEYTFHLYLIPDQCSIR	1260
QY	1261	K-----ELFCYRSRPGDQFL	1276
Db	1261	KALIDLEMKQFVARIHKKPPLTPLYNGCRVTSGSGRDAGNTPQELTECYRSPGEGDQFL	1320
QY	1277	SEFVYGHLSGSGIRLQVKKKDETLWEALVKPQDLMPATLLIPRAIAVSPIDAPQLLH	1336
Db	1321	SEFVYGHLSGSGIRLQVKKKDETLWEALVKPQDLMPATLLIPRAIAVSPIDAPQLLH	1380
QY	1337	FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAEVTRPSQWRKLFSLSQSWDRKC	1396
Db	1381	FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAEVTRPSQWRKLFSLSQSWDRKC	1440
QY	1397	KDGLYQALKETHPHLIMELMEKSKKGLPLSS	1429
Db	1441	KDGLYQALKETHPHLIMELMEKSKKGLPLSS	1473
RESULT 9			
ABG78455	ABG78455 standard; Protein; 1429 AA.		
AC	ABG78455;		
DT	15-NOV-2002 (first entry)		
DE	Human caspase recruitment protein 7 protein.		
XX	Human; human leucine-rich repeat small intestine I; HLRRSII; asthma;		
KW	proliferative disorder; gastrointestinal disorder; renal disorder;		
KM	neural disorder; reproductive disorder; calcium regulation; apoptosis;		
KW	immune system; anaemia; human immune deficiency virus; HIV; cancer;		
KM	blood coagulation disorder; autoimmune disorder; allergic reaction;		
KW	inflammatory condition; cardiovascular disorder; ischaemia;		
KM	neurological disorder; infectious disease; cytokine production;		
KW	expressed sequence tag; BGT.		
XX	Homo sapiens.		
OS	MO200261086-A2.		
PN	08-AUG-2002.		
PD	20-DEC-2001; 2001MO-US49739.		
FP	22-DEC-2000; 2000US-257774P.		
XX			

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Feder J, Ramanathan C, Mintier G;
 PI WPI; 2002-619252/66.
 XX
 XX
 PT New isolated nucleic acid molecules encoding HLRSSII polypeptides, or
 PT their fragments and homologues, useful for preventing, treating and
 PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
 PT or renal disorders
 PS Disclosure; Figure 2; 336pp; English.
 XX
 XX The invention relates to isolated nucleic acid molecules (1) encoding
 CC human leucine-rich repeat small intestine I (HLRSSII) polypeptides.
 CC The nucleic acid molecules and polypeptides are useful for preventing,
 CC treating and ameliorating medical conditions, such as proliferative,
 CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
 CC related to aberrant calcium regulation or apoptosis modulation, either
 CC directly or indirectly. They are also useful for treating, preventing
 CC and/or diagnosing diseases, disorders and/or conditions of: immune system
 CC by activating or inhibiting the proliferation, differentiation, or
 CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
 CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
 CC virus (HIV) infection, HTLV-BLV infection; blood coagulation disorders,
 CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
 CC myasthenia gravis; asthma or allergic reactions; inflammatory
 CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
 CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
 CC ischaemia; aneurysms; neurological disorders, e.g. Alzheimer's disease,
 CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
 CC pneumonia, or viral, bacterial, and fungal infections. The HLRSSII
 CC polypeptides are useful for modulating cytokine production, antigen
 CC presentation, or other processes such as boosting immune responses.
 CC ABG7454-ABG7484 represent HLRSSII amino acid sequences and related
 CC amino acid sequences of the invention.
 XX
 XX Sequence 1429 AA;
 Query Match 99.4%; Score 7488; DB 23; Length 1429;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MAGGAGRLACVLEFLKKEELKFPOLLANKASRSSGPTPAOPEKTSMEVASYLVQA 60
 DB 1 MAGGAGRLACVLEFLKKEELKFPOLLANKASRSSGPTPAOPEKTSMEVASYLVQA 60
 QY 61 YGEORAMDLALHTWEGWGLRSLCAQAQEGAGHSPFPYSPSEPHLGSPQSTAVLMPV 120
 DB 61 YGEORAMDLALHTWEGWGLRSLCAQAQEGAGHSPFPYSPSEPHLGSPQSTAVLMPV 120
 QY 121 IHELPAGCTGGSSRRVRLQPLDTSRRRRREISASLLYQALPSSPDHESPQSPNAPTST 180
 DB 121 IHELPAGCTGGSSRRVRLQPLDTSRRRRREISASLLYQALPSSPDHESPQSPNAPTST 180
 QY 181 AUVGSMGSPPOSLAPRQAPAGTQWPLDTSGLYYTEIERERRESEKRPMAAVGT 240
 DB 181 AUVGSMGSPPOSLAPRQAPAGTQWPLDTSGLYYTEIERERRESEKRPMAAVGT 240
 QY 241 PPOAHLISLOPHHHPWESVRESLCSWPMKQEDFNQKFTQLLLQRPHPASQPLVRSW 300
 DB 241 PPOAHLISLOPHHHPWESVRESLCSWPMKQEDFNQKFTQLLLQRPHPASQPLVRSW 300
 QY 301 PDVVEENRGLIIRDLFGPGLDQEPRIYILQAGAGIGKSTLARQYKAMGRGQLYGDR 360
 DB 301 PDVVEENRGLIIRDLFGPGLDQEPRIYILQAGAGIGKSTLARQYKAMGRGQLYGDR 360
 QY 361 FOHVFVSSCELSQSKVSLAELIGKGTATPAPIROIILSRPRLFIILGNDPBGVILQ 420
 DB 361 FOHVFVSSCELSQSKVSLAELIGKGTATPAPIROIILSRPRLFIILGNDPBGVILQ 420
 QY 421 EPSESELCHMSQOPADALGSLIGKTIILPEASFLITARTTALONILPSLEQARWVVLG 480

DB 421 EPSESELCHMSQOPADALGSLIGKTIILPEASFLITARTTALONILPSLEQARWVVLG 480
 QY 481 FSESSRKEYFYRYFTDERQAIARFVLKSNKEIMALCLVPWVSWLACTCLMOQKREKYL 540
 DB 481 FSESSRKEYFYRYFTDERQAIARFVLKSNKEIMALCLVPWVSWLACTCLMOQKREKYL 540
 QY 541 TLTSKTTTLCHLYLAQALQAOPLGQPLDLSIAEGIWOKKTLFSPDDLRHGDGAI 600
 DB 541 TLTSKTTTLCHLYLAQALQAOPLGQPLDLSIAEGIWOKKTLFSPDDLRHGDGAI 600
 QY 601 ISTFLMGILQEHPIPLSYSPFHLCPQEPFAAMSYLEDEKRGKSNICIDLEKTLFAY 660
 DB 601 ISTFLMGILQEHPIPLSYSPFHLCPQEPFAAMSYLEDEKRGKSNICIDLEKTLFAY 660
 QY 661 GHHGFGASTTRPLGLSDEGBREKENI FHCRLSGRNLMQWVPSLOLLLOPHSLSLH 720
 DB 661 GHHGFGASTTRPLGLSDEGBREKENI FHCRLSGRNLMQWVPSLOLLLOPHSLSLH 720
 QY 721 CLYERPKTFLQVMAHFEEMQCVETDMLVCFICPSRHVKLOLIEGRQHSITS 780
 DB 721 CLYERPKTFLQVMAHFEEMQCVETDMLVCFICPSRHVKLOLIEGRQHSITS 780
 QY 781 PTWVVLFRWVPTDAVWQILFVLKVTNMLKEILDSGNSLSHSAVSKLCTKLRRPCLLE 840
 DB 781 PTWVVLFRWVPTDAVWQILFVLKVTNMLKEILDSGNSLSHSAVSKLCTKLRRPCLLE 840
 QY 841 TLRACGGLTAEBCDCLAFGLRANQTLTELDSFNVLTDAAGXHCQRLRPSCKLQRIQ 900
 DB 841 TLRACGGLTAEBCDCLAFGLRANQTLTELDSFNVLTDAAGXHCQRLRPSCKLQRIQ 900
 QY 901 LVSCGILTSCCDDLASVLSASPSIKELDIQONNLDVGRILLCEGRHRAKTLIRGLDQ 960
 DB 901 LVSCGILTSCCDDLASVLSASPSIKELDIQONNLDVGRILLCEGRHRAKTLIRGLDQ 960
 QY 961 TILSDERQELALBEQEKQQLIFSRKRPVWPTFEGDITGEMSNSTSLKQRLSERA 1020
 DB 961 TILSDERQELALBEQEKQQLIFSRKRPVWPTFEGDITGEMSNSTSLKQRLSERA 1020
 QY 1021 ASHVAQANIKLIDVSKIPFIABIAEBSPEVVPVELLCVPSPASQCDLTKPLGTDDDPW 1080
 DB 1021 ASHVAQANIKLIDVSKIPFIABIAEBSPEVVPVELLCVPSPASQCDLTKPLGTDDDPW 1080
 QY 1081 GPTGPVATEVVDKKNLYVHFPPVAGSVWPTGCLCFVREAVTVIEECWMDQFLGEIN 1140
 DB 1081 GPTGPVATEVVDKKNLYVHFPPVAGSVWPTGCLCFVREAVTVIEECWMDQFLGEIN 1140
 QY 1141 PHSWVWAGPLLDIKAPGAVEAVHLPHFVALQGGHVDTSLPQVAFKESGMLLEKPAV 1200
 DB 1141 PHSWVWAGPLLDIKAPGAVEAVHLPHFVALQGGHVDTSLPQVAFKESGMLLEKPAV 1200
 QY 1201 ELHHIVLENPSPPLGAVLLKMTIHNAIRTPVTSVVLVYHRVHPEEYTFHLIIPSCSTR 1260
 DB 1201 ELHHIVLENPSPPLGAVLLKMTIHNAIRTPVTSVVLVYHRVHPEEYTFHLIIPSCSTR 1260
 QY 1261 KELELCYRSPGSDOLPSEFYVHLSGIRLOYKDKDETLWEALVKPDIMPATTLIP 1320
 DB 1261 KELELCYRSPGSDOLPSEFYVHLSGIRLOYKDKDETLWEALVKPDIMPATTLIP 1320
 QY 1321 ARIAVSPPLDAPQLHFVDQYREOLARVTSVEVVLIDKLHGQVLSQEQERYVLAENTRPS 1380
 DB 1321 ARIAVSPPLDAPQLHFVDQYREOLARVTSVEVVLIDKLHGQVLSQEQERYVLAENTRPS 1380
 QY 1381 QWRKLFSLQSQWDRCKDGLYQALKXTHPHLIMELMEKSKKGLPLSS 1429
 DB 1381 QWRKLFSLQSQWDRCKDGLYQALKXTHPHLIMELMEKSKKGLPLSS 1429
 RESULT 10
 AA017855 standard; Protein; 1429 AA.
 ID AA017855
 XX
 AC AA017855;

XX 20-AUG-2002 (first entry)

XX Pyrin domain containing protein NALP1-bs.

XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 XX antiarteriosclerotic; antipsoriatic; antibacterial; vincidic;
 XX neuroprotective; antiarthritic; antirheumatic; antiaesthetic;
 XX neurotropic; osteopathic; nootropic; intracellular signal transduction;
 XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 XX osteoarthritis; glomerulonephritis.

XX Unidentified.

XX WO200240668-A2.

XX 23-MAY-2002.

XX 30-OCT-2001; 2001WO-EP12545.

XX 15-NOV-2000; 2000DE-1056687.
 XX 30-NOV-2000; 2000DE-1059595.

XX (APOT-) APOTEC RES & DEV LTD.

XX Tschopp J, Martinon F;
 XX MPI; 2002-427093/45.
 XX N-PSDB; AAL47127.

XX New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies

XX Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention.

XX Sequence 1429 AA;

XX Query Match 99.4%; Score 7488; DB 23; Length 1429;
 XX Best Local Similarity 99.4%; Pred. No. 0;
 XX Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASYVAQ 60
 DB 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASYVAQ 60

QY 61 YGEORANDLALHTWQGLSLCAQOAGAGHSPSPYSPSEPHLGSPOPTSTAVLMPW 120
 DB 61 YGEORANDLALHTWQGLSLCAQOAGAGHSPSPYSPSEPHLGSPOPTSTAVLMPW 120

QY 121 IHEHPACCTGSSERVVRLPDTSGRRWREISASHLYQALPSSPDHSPSGESSTNAPTST 180
 DB 121 IHEHPACCTGSSERVVRLPDTSGRRWREISASHLYQALPSSPDHSPSGESSTNAPTST 180

QY 121 IHEHPACCTGSSERVVRLPDTSGRRWREISASHLYQALPSSPDHSPSGESSTNAPTST 180
 DB 121 IHEHPACCTGSSERVVRLPDTSGRRWREISASHLYQALPSSPDHSPSGESSTNAPTST 180

QY 181 AVLGSMSPPOPSLAPREOAPGTOMPLDSTGYTYEIREREREKSEKGPMAAVVGT 240
 DB 181 AVLGSMSPPOPSLAPREOAPGTOMPLDSTGYTYEIREREREKSEKGPMAAVVGT 240

QY 241 PPOAHTSLQPHHHEWSPVRESLSTWPKNEDFNOKFTOLLQPHPRSQDPLVRSW 300
 DB 241 PPOAHTSLQPHHHEWSPVRESLSTWPKNEDFNOKFTOLLQPHPRSQDPLVRSW 300

QY 301 PDVEENRGHLIRIDLFGBGLDQEPRIYVILQGAAGIKSTLAROYKAMGRGOLYGR 360
 DB 301 PDVEENRGHLIRIDLFGBGLDQEPRIYVILQGAAGIKSTLAROYKAMGRGOLYGR 360

QY 361 FOHVFPSCRELAOSKVSJLAELIGKGTATPAPIROILSRPERLLTLDGVDSPGWLQ 420
 DB 361 FOHVFPSCRELAOSKVSJLAELIGKGTATPAPIROILSRPERLLTLDGVDSPGWLQ 420

QY 421 EPSEELCLHMSQOPADALLGSLGKTIIEBASFLITARTTALQNLPSLEQARWEVLG 480
 DB 421 EPSEELCLHMSQOPADALLGSLGKTIIEBASFLITARTTALQNLPSLEQARWEVLG 480

QY 481 FSSSRKREYFRYFTDROAIRAPRLVSKNKMALCLVWVSVLACTCLMOQKREKL 540
 DB 481 FSSSRKREYFRYFTDROAIRAPRLVSKNKMALCLVWVSVLACTCLMOQKREKL 540

QY 541 TLNKTITTLCLMYLAQALQAPLGPOLRDLCSLAAGIWMOKTLFSPDDLRRKGLGAI 600
 DB 541 TLNKTITTLCLMYLAQALQAPLGPOLRDLCSLAAGIWMOKTLFSPDDLRRKGLGAI 600

QY 601 ISTFLKXGILQEHPIPLSYSPFHLCFQEPFAAMSYYLEDEKRGKHSNCIIDEKTLAY 660
 DB 601 ISTFLKXGILQEHPIPLSYSPFHLCFQEPFAAMSYYLEDEKRGKHSNCIIDEKTLAY 660

QY 661 GIGHLFGASTTRPLGLSDGEREMENIFCHRLSQRNLMOWPSTQLLOPHSLSLH 720
 DB 661 GIGHLFGASTTRPLGLSDGEREMENIFCHRLSQRNLMOWPSTQLLOPHSLSLH 720

QY 721 CLVETRRKFTLTQWMAFEEMGMCVETDMELVCTFCIKFSRHVKQLIIEGRORSTWS 780
 DB 721 CLVETRRKFTLTQWMAFEEMGMCVETDMELVCTFCIKFSRHVKQLIIEGRORSTWS 780

QY 781 PTWVLEFRWVPYTDVWQILFSVLKVRNLKELDLGSNSLSHSAVKSLCKTLRRPRCLLE 840
 DB 781 PTWVLEFRWVPYTDVWQILFSVLKVRNLKELDLGSNSLSHSAVKSLCKTLRRPRCLLE 840

QY 841 TLRLAAGGLAEDCKOLAFGLRANQTLTELDLSPNYTDAKALCORLROPSCKLQRLQ 900
 DB 841 TLRLAAGGLAEDCKOLAFGLRANQTLTELDLSPNYTDAKALCORLROPSCKLQRLQ 900

QY 901 LVSCGLTSDCCODLAVSLASPSLKELDLQONNLDVGVALLCEGLHPACKLIRLGLDQ 960
 DB 901 LVSCGLTSDCCODLAVSLASPSLKELDLQONNLDVGVALLCEGLHPACKLIRLGLDQ 960

QY 961 TTSLSDEMRQELRALBOEKPOLLIFSRKPSVMTPTBGLDTGEMSNSTSLKROFLSERA 1020
 DB 961 TTSLSDEMRQELRALBOEKPOLLIFSRKPSVMTPTBGLDTGEMSNSTSLKROFLSERA 1020

QY 1021 ASHVAQANLKLIDVSKTFPIAETAEBSSEPVVPELLCVSPASQGLLHFKPLGTDDDFW 1080
 DB 1021 ASHVAQANLKLIDVSKTFPIAETAEBSSEPVVPELLCVSPASQGLLHFKPLGTDDDFW 1080

QY 1081 GPTGPVATEVVDKESKILYRHPVAGSYRMPNTGLCFVMEBAVTVVEIFCWMDQFLGIN 1140
 DB 1081 GPTGPVATEVVDKESKILYRHPVAGSYRMPNTGLCFVMEBAVTVVEIFCWMDQFLGIN 1140

QY 1141 PSHSMVAVGELLDIKAEPGAVEAVLHPFVALQGHVDTSLFCVAHFKESGMLLEKPARV 1200
 DB 1141 PSHSMVAVGELLDIKAEPGAVEAVLHPFVALQGHVDTSLFCVAHFKESGMLLEKPARV 1200

QY 1201 ELKHIVLENPSFPLGVLKMINALFIVTSTVLLYHRLHPEEVTHLYLPSDCSIR 1260
 DB 1201 ELKHIVLENPSFPLGVLKMINALFIVTSTVLLYHRLHPEEVTHLYLPSDCSIR 1260

QY 1261 KELELCYRSGEDOLFSEFVYVGHLSGIRLQVQDKODETLVWEALVKGDMAPATTLIP 1320
 DB 1261 KELELCYRSGEDOLFSEFVYVGHLSGIRLQVQDKODETLVWEALVKGDMAPATTLIP 1320

QY 1321 ARLAVPSPLDAPOLLHFVDQYRROLARVTSVEVLDKHLGQVLSQGYERVLAEENTRPS 1380
 DB 1321 ARLAVPSPLDAPOLLHFVDQYRROLARVTSVEVLDKHLGQVLSQGYERVLAEENTRPS 1380

QY 1381 QMRKLPSLSQSWDRKCKDGLYQALKETHPHLIMELWEXSKKGLLPSS 1429

DB 1381 QMRKLFSLISQSWDRCKDGLYQALKEPHLIMELWEEKSKGGLPLSS 1429

RESULT 11

AAE06758
ID AAE06758 standard; Protein; 1473 AA.

AAE06758;

16-OCT-2001 (first entry)

Human G-protein coupled receptor-8 (GCR8-8) protein.

Human; G-protein coupled receptor-8; GCR8-8; cytosolic; hepatotropic; virocidic; antiinflammatory; anticonvulsant; antiemetic; neuroprotective; neurotrophic; cerebroprotective; hypotensive; tranquilizer; vulnerary; ophthalmological; cell proliferative disorder; actinic keratosis; anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease; Huntington's disease; Parkinson's disease; cardiovascular disorder; epilepsy; hypertension; varicose vein; vasculitis; dyspepsia; anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder; Addison's disease; Crohn's disease; acquired immune deficiency syndrome; AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity; osteoporosis; transgenic animal; gene therapy.

Homo sapiens.

Key Location/Qualifiers

Domain 1216..1237

Binding-site 334..341

/label=ATP/GTP-binding_site

/note="P-loop"

MO200157085-A2.

09-AUG-2001.

01-FEB-2001; 2001WO-US03455.

02-FEB-2000; 2000US-0180093.

11-FEB-2000; 2000US-0182045.

(INCY-) INCYTE GENOMICS INC.

Baughn MR, Au-Young J, Yue H;

WPI; 2001-48869/53.

N-PSDB; AAD12951.

Novel isolated human G-protein coupled receptor useful for diagnosing,

preventing and treating cell proliferative, neurological,

cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic

disorders -

Claim 1; Page 114-117; 138pp; English.

The present sequence is human G-protein coupled receptor-8 (GCR8-8) protein. The present invention relates to GCR8 protein and nucleic acids encoding them. GCR8 protein, its agonist or antagonist are useful for treating diseases or conditions associated with decreased expression or overexpression of functional GCR8 in a patient, where the disorder is selected from cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and cancer, neurological disorders such as epilepsy, stroke, Alzheimer's disease, Huntington's disease, Parkinson's disease, cardiovascular disorders such as hypertension, vasculitis, varicose veins, gastro-intestinal disorders such as dyspepsia, dyspepsia, anorexia, nausea, pancreatitis, autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic

infections, trauma and metabolic disorders such as diabetes, obesity, osteoporosis. GCR8 proteins and their cDNAs are used to assess the effects of exogenous compounds on the expression of GCR8 sequences. GCR8 cDNA is useful to create knock in humanised animals (p19s) or transgenic animals (mice or rats) to model human disease, for therapeutic or diagnostic purposes, for somatic or germ-line gene therapy, to generate hybridisation probes useful in mapping the CC naturally occurring genomic sequence, and in molecular biological techniques.

Sequence 1473 AA;

Query Match 99.3%; Score 7484; DB 22; Length 1473;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 1426; Conservative 1; Mismatches 2; Indels 44; Gaps 1;

1 MAGGAMGRACYLEEFLKKEBELKEPOLLANKHSSSSGEPAPQEKTSGMVAATLVQ 60
1 MAGGAMGRACYLEEFLKKEBELKEPOLLANKHSSSSGEPAPQEKTSGMVAATLVQ 60
61 YGEORAMDLATMTWQMGRLSCAQAOGAGHSPSPSPSPSPSPSPSPSPSPSPSPSP 120
61 YGEORAMDLATMTWQMGRLSCAQAOGAGHSPSPSPSPSPSPSPSPSPSPSPSPSP 120
121 IHFLPAGCTQGSERVLKPLPTSGRRWRREISALLYQALPSPSPSPSPSPSPSPSP 180
121 IHFLPAGCTQGSERVLKPLPTSGRRWRREISALLYQALPSPSPSPSPSPSPSPSP 180
181 AVLGSMGSPPOPSLAPREOAPGTOMPLDETSGIYYTEREEREKSEKGRPPMAVGT 240
181 AVLGSMGSPPOPSLAPREOAPGTOMPLDETSGIYYTEREEREKSEKGRPPMAVGT 240
241 PQAATSLQPHHHPWSPVRESLCSWPWKNDENQKFTOLLQRPSPSPSPSPSPSPSP 300
241 PQAATSLQPHHHPWSPVRESLCSWPWKNDENQKFTOLLQRPSPSPSPSPSPSPSP 300
301 PDYVEENRGLHLEIFDLPFGGLDTPERIVILQGAAGIGKTLARQVEMWGRGOLYGR 360
301 PDYVEENRGLHLEIFDLPFGGLDTPERIVILQGAAGIGKTLARQVEMWGRGOLYGR 360
361 FOHVFPYSCRELAQSKVSLAELIGKGTATPAPRQILSRPERLLFLDGDPEPMVQ 420
361 FOHVFPYSCRELAQSKVSLAELIGKGTATPAPRQILSRPERLLFLDGDPEPMVQ 420
421 EPSSSELCLHWSQPOPADNLGSLGKTLIPASFLITATTALQNLIPLEQARVAVLG 480
421 EPSSSELCLHWSQPOPADNLGSLGKTLIPASFLITATTALQNLIPLEQARVAVLG 480
481 PESSRKEFFRYFPTDERQAIAPFLVYSNKLMLCLVPWVSWLACTLMQMKRKEKL 540
481 PESSRKEFFRYFPTDERQAIAPFLVYSNKLMLCLVPWVSWLACTLMQMKRKEKL 540
541 TTSKTTTLCIHYAQAOLQAPLQRLDLSLAEGIWQKTLFSPDDLRKHGIDGAI 600
541 TTSKTTTLCIHYAQAOLQAPLQRLDLSLAEGIWQKTLFSPDDLRKHGIDGAI 600
601 ISTFLKMGILQHPILPLSYFTHLCFOEFPFAMSVYLEDKRGKXSNCIIDLEKLEY 660
601 ISTFLKMGILQHPILPLSYFTHLCFOEFPFAMSVYLEDKRGKXSNCIIDLEKLEY 660
661 GHHGFGASTTFLGLISDEGERMENTFHRLOSGRMLQWVSLQILLPHSIESAH 720
661 GHHGFGASTTFLGLISDEGERMENTFHRLOSGRMLQWVSLQILLPHSIESAH 720
721 CLETFRNKTLFLQWMAHFEEMQCVETDMLVCFICIFSRHVYKLQLEBGRHSTWS 780
721 CLETFRNKTLFLQWMAHFEEMQCVETDMLVCFICIFSRHVYKLQLEBGRHSTWS 780
781 PTMVTLEFRVPTDAVWQILSVLKATRNKLKLDLSGNSLSAYKSLKTLRRPCLLE 840
781 PTMVTLEFRVPTDAVWQILSVLKATRNKLKLDLSGNSLSAYKSLKTLRRPCLLE 840
841 TRLAGCGTLADCKDLAFGLRANQTLTFLDLSPVLTDAKAKHLCQRLRQPSCKLQRLQ 900


```

QY 481 FSSSRKRYFYRFTDERQAIKRAFLVKSNEKIMALCLVWVSWLACTCIQMOKREKL 540
DB 481 FSSSRKRYFYRFTDERQAIKRAFLVKSNEKIMALCLVWVSWLACTCIQMOKREKL 540
QY 541 TLTSKTTTTLCLHYLAQALQAPOLGPDRLDGLSAAEGIMQKXTLSPDDLKRGHLDGAI 600
DB 541 TLTSKTTTTLCLHYLAQALQAPOLGPDRLDGLSAAEGIMQKXTLSPDDLKRGHLDGAI 600
QY 601 ISTFLKMGIIQEHPIPLSYSFHLCFOEPFAMSVYLEDEKGRKSNCCIIDEKTLKAV 660
DB 601 ISTFLKMGIIQEHPIPLSYSFHLCFOEPFAMSVYLEDEKGRKSNCCIIDEKTLKAV 660
QY 661 GIGLFGASTTRFLGLSDEGEREMENIFHCRLSGRNLMQWPSIQLLQPHSLSLH 720
DB 661 GIGLFGASTTRFLGLSDEGEREMENIFHCRLSGRNLMQWPSIQLLQPHSLSLH 720
QY 721 CLVETRKFTFLTOVMAFEEMGMCVETDMLVCTFCIKFSRHVKQLQLEGRQHSHTWS 780
DB 721 CLVETRKFTFLTOVMAFEEMGMCVETDMLVCTFCIKFSRHVKQLQLEGRQHSHTWS 780
QY 781 PTWVVLFRWVPTDVAWYQILFSVLKVTNKLKELDSGNSLSHSAYVSLCKTLRRPCLLE 840
DB 781 PTWVVLFRWVPTDVAWYQILFSVLKVTNKLKELDSGNSLSHSAYVSLCKTLRRPCLLE 840
QY 841 TFLRAGCGLTABDCKDLAFGLRANQITLEDLSFNVLTDGAKHLCORLQPSCKLQRLQ 900
DB 841 TFLRAGCGLTABDCKDLAFGLRANQITLEDLSFNVLTDGAKHLCORLQPSCKLQRLQ 900
QY 901 LVSCGGLTSDCCODLASVLSAPSLKELDLQNNLDVGVALLCEGLRHPRCKLIRLGLDQ 960
DB 901 LVSCGGLTSDCCODLASVLSAPSLKELDLQNNLDVGVALLCEGLRHPRCKLIRLGLDQ 960
QY 961 TFLSDENRQELRALDEKPOLLIFSRKRPVMTPTBGLDTEGMSNSTSSLKQRLGSERA 1020
DB 961 TFLSDENRQELRALDEKPOLLIFSRKRPVMTPTBGLDTEGMSNSTSSLKQRLGSERA 1020
QY 1021 ASHVAQNLKLDVSKIFPIAETIAESSPEVVPBELLCVSPASQGLHTRKPLGTDDEFW 1080
DB 1021 ASHVAQNLKLDVSKIFPIAETIAESSPEVVPBELLCVSPASQGLHTRKPLGTDDEFW 1080
QY 1081 GPTGPVATEVVDKKNLYRHVPVAGSYRWPNGLCFVMBEATVVEIEFCVWQDFLGEIN 1140
DB 1081 GPTGPVATEVVDKKNLYRHVPVAGSYRWPNGLCFVMBEATVVEIEFCVWQDFLGEIN 1140
QY 1141 PQSHMWVAGPLLDIKAEKPAVEAVLHPFVALQGHVDTSLFQVAFKESGMLLEKPARV 1200
DB 1141 PQSHMWVAGPLLDIKAEKPAVEAVLHPFVALQGHVDTSLFQVAFKESGMLLEKPARV 1200
QY 1201 ELHHIVLENSFSPFLGVLKMIHNALEFIPVTSVLLYHNVHPEVTFFHLYLIPSDCSIR 1260
DB 1201 ELHHIVLENSFSPFLGVLKMIHNALEFIPVTSVLLYHNVHPEVTFFHLYLIPSDCSIR 1260
QY 1261 KELELCYRSPGEBDLFSEFVYVGHLSGIRLOVKDKDETVMELVYRGDLMPATTLIP 1320
DB 1261 KELELCYRSPGEBDLFSEFVYVGHLSGIRLOVKDKDETVMELVYRGDLMPATTLIP 1320
QY 1321 ARIAVPSPLDAPOLHAFVDQIREQLIARVTSVEVLDKHLGOVLQSOBYERVLAEENTRPS 1380
DB 1321 ARIAVPSPLDAPOLHAFVDQIREQLIARVTSVEVLDKHLGOVLQSOBYERVLAEENTRPS 1380
QY 1381 QMRKLPFLSOSWDRCKDGLYQALKETHPHLIMELWKSGLPLSS 1429
DB 1381 QMRKLPFLSOSWDRCKDGLYQALKETHPHLIMELWKSGLPLSS 1429
QY 1429 QMRKLPFLSOSWDRCKDGLYQALKETHPHLIMELWKSGLPLSS 1397
DB 1429 QMRKLPFLSOSWDRCKDGLYQALKETHPHLIMELWKSGLPLSS 1397

```

RESULT 13

AAV72671 standard; Protein; 1442 AA.

XX AAV72671;

XX 31-MAY-2001 (first entry)

XX

```

DE Human NB-ARC and CARD containing protein (NAC) delta isoform.
XX
XX Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
XX caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
XX cysteine protease; apoptosis; cytokine production;
XX cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
XX fibrosis; arthritis; cancer; adenocarcinoma; leukemia.
XX
XX Homo sapiens.
OS
XX
XX Key
XX Region
XX 1..956
XX /note="Corresponds to 1-956 residues of human NAC
XX beta isoform (AAV72669)"
XX Region
XX /note="Corresponds to 988-1473 residues of human NAC
XX beta isoform (AAV72669)"
XX
XX WO200116170-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-US24152.
XX
XX 01-SEP-1999; 99US-0388221.
XX
XX (BURN-) BURHAM INST.
XX
XX Reed JC;
XX
XX WPI, 2001-183258/18.
XX
XX N-PSDB; AAD02762.
XX
XX Novel nucleic acid encoding NB-ARC and caspase associated recruitment
XX domains; used to produce polypeptides for screening for modulators of
XX apoptosis
XX
XX PT
XX Claim 15; Page 154-158; 184pp; English.
XX
XX PS
XX
XX The present sequence is a human NB-ARC and CARD containing protein
XX (NAC) delta isoform. NAC delta isoform represents the NAC splice
XX variant in which one of the splice region is absent in the translated
XX polypeptide. NAC protein comprises a nucleotide binding (NB) domain
XX (also referred as NB-ARC domain), a caspase-associated recruitment
XX domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
XX aspartyl proteases, are principal effectors of apoptosis. CARD containing
XX NAC proteins are used for screening modulators that modulates apoptosis,
XX cytokine production, cytokine receptor signalling and other cellular
XX processes. NAC can act as an immunogen for the production of polyclonal
XX and monoclonal antibodies. It can also be used to diagnose and treat
XX inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
XX pathologies such as adenocarcinomas and leukemias.
XX
XX SO
XX Sequence 1442 AA;
XX
XX Query Match 97.2%; Score 7325.5; DB 22; Length 1442;
XX Best Local Similarity 94.9%; Pred. No. 0;
XX Matches 1398; Conservative 0; Mismatches 0; Indels 75; Gaps 2;
XX
XX 1 MAGGAGRLACTYELFKKEELKPOLLLANKASRSSSETPAQPKTSGMEVASYLVNQ 60
XX 1 MAGGAGRLACTYELFKKEELKPOLLLANKASRSSSETPAQPKTSGMEVASYLVNQ 60
DB 61 YGEQRAMDALHTWBMGLRSLCAQAGBAGHSPPFPSPSEPHLGSPOQPTSTAVLMPW 120
DB 61 YGEQRAMDALHTWBMGLRSLCAQAGBAGHSPPFPSPSEPHLGSPOQPTSTAVLMPW 120
QY 121 IHELPAGCTQSGSRRLVROLPTSGRRMEIASLLYQALPSPDHESSESQSPNAPTST 180
DB 121 IHELPAGCTQSGSRRLVROLPTSGRRMEIASLLYQALPSPDHESSESQSPNAPTST 180
QY 181 AVTGSWSPPOPSLAREQAPGTOWPLDETSGIYYTEIRERRESEKGRPPMAVVG 240
DB 181 AVTGSWSPPOPSLAREQAPGTOWPLDETSGIYYTEIRERRESEKGRPPMAVVG 240

```

Db 181 AVLGWSGPPQPSLADREQAPGTQWPLDSTGTYTEIRERERSEKGRPEMAAVGT 240
 Qy 241 PPOAHSLQPHHHPBWSVRESLCTWPKNEDPNQFTOLLQPHRPSQDPLYKRSW 300
 Db 241 PPOAHSLQPHHHPBWSVRESLCTWPKNEDPNQFTOLLQPHRPSQDPLYKRSW 300
 Qy 301 PDVYENRGLIIRDLFGPGLDTPQPRVITLQAGAGIKSTLAROYKEMGQLYGDR 360
 Db 301 PDVYENRGLIIRDLFGPGLDTPQPRVITLQAGAGIKSTLAROYKEMGQLYGDR 360
 Qy 361 FQHVFFSCHELAQSKVSLAEILGKDTATPARPIROILSRPERLLFIIDVDEPGVNLQ 420
 Db 361 FQHVFFSCHELAQSKVSLAEILGKDTATPARPIROILSRPERLLFIIDVDEPGVNLQ 420
 Qy 421 EBSSELCIHWSPQPADALIGSLGKTIPEASFLITARTTAQNLIPSEQARWVYLQ 480
 Db 421 EBSSELCIHWSPQPADALIGSLGKTIPEASFLITARTTAQNLIPSEQARWVYLQ 480
 Qy 481 FSSSRKREYFRFTDBROAIRAFRLVKSNEKLMALCLVPWVSMIACGLMOQMKREKL 540
 Db 481 FSSSRKREYFRFTDBROAIRAFRLVKSNEKLMALCLVPWVSMIACGLMOQMKREKL 540
 Qy 541 TLTSKTTTTLCLHYLAQALQAPLQPLQDLCSIAEGIQKXTLPSDDLRKHGLDGA 600
 Db 541 TLTSKTTTTLCLHYLAQALQAPLQPLQDLCSIAEGIQKXTLPSDDLRKHGLDGA 600
 Qy 601 ISTFLKMGIIQHPPIPLSYFHLCFQEPFPAAMSYLEDEKRGKHSNCIIDLEKTLAY 660
 Db 601 ISTFLKMGIIQHPPIPLSYFHLCFQEPFPAAMSYLEDEKRGKHSNCIIDLEKTLAY 660
 Qy 661 GINGLGASTTRFLGLISDEGEREMENIFHCISQSRNIMQWVPSIQLLOPHSLESH 720
 Db 661 GINGLGASTTRFLGLISDEGEREMENIFHCISQSRNIMQWVPSIQLLOPHSLESH 720
 Qy 721 CLVETRNKFTLTQVMAHFEEMKGCVEITMELLVCTFCIKFSRYVKKQLIEGRQHSW 780
 Db 721 CLVETRNKFTLTQVMAHFEEMKGCVEITMELLVCTFCIKFSRYVKKQLIEGRQHSW 780
 Qy 781 PTWVVLFRWVPTVDAYWQIILFSVLKVTNRNKELDLSGNSLSHSAVSLCCTLRPRCLLE 840
 Db 781 PTWVVLFRWVPTVDAYWQIILFSVLKVTNRNKELDLSGNSLSHSAVSLCCTLRPRCLLE 840
 Qy 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDSLNNVLTDAAGHLCORLQOPSCKLQ 900
 Db 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDSLNNVLTDAAGHLCORLQOPSCKLQ 900
 Qy 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 Db 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 Qy 961 TLLSDENRQELRALBOEKPOLLIFSRRKPSVMTPTBGLDTGEMSNSTSSLRQRLSERA 1020
 Db 961 TLLSDENRQELRALBOEKPOLLIFSRRKPSVMTPTBGLDTGEMSNSTSSLRQRLSERA 1020
 Qy 990 ASHVAQNLKLLDVSKIFPIAETAEBSSEPVVPELLCVSPASQGLHKKPGLTDDDFW 1080
 Db 990 ASHVAQNLKLLDVSKIFPIAETAEBSSEPVVPELLCVSPASQGLHKKPGLTDDDFW 1080
 Qy 1081 GPTGPVATEVVDKELNLYRHFPAVAGSYRNPNTGLCFVMEBEAVTELEFCVMDQFLGEIN 1140
 Db 1081 GPTGPVATEVVDKELNLYRHFPAVAGSYRNPNTGLCFVMEBEAVTELEFCVMDQFLGEIN 1140
 Qy 1141 PHSWVAVAGPLLDIKABPGAVEAHLPHFVALQGHVDTSLFQMAHFKESGMLLEKPARV 1200
 Db 1141 PHSWVAVAGPLLDIKABPGAVEAHLPHFVALQGHVDTSLFQMAHFKESGMLLEKPARV 1200
 Qy 1201 ELHHIVLENSFSPGLIKMHNALRFIPVTSVLLVYHVPBEVPHYLPIPSDSIR 1260
 Db 1201 ELHHIVLENSFSPGLIKMHNALRFIPVTSVLLVYHVPBEVPHYLPIPSDSIR 1260
 Qy 1261 -----KELSLCYRSPGSDQLF 1276
 Db 1230 KALDLEMKQFVRIHKKPPLTPLYMGCRTYVSSGSGMLEILPKELSLCYRSPGSDQLF 1289

Qy 1277 SEFYVGHLSGSRIRLQYKDKKDETLVWEALVKBGDLMPATLLIPPAIAYSPIDAPOLH 1336
 Db 1290 SEFYVGHLSGSRIRLQYKDKKDETLVWEALVKBGDLMPATLLIPPAIAYSPIDAPOLH 1349
 Qy 1337 FVDQYREQLIARTSVAVVLDKLGQVLSQEQERYVLAENTRPSQMRKLFSLSQSMDRC 1396
 Db 1350 FVDQYREQLIARTSVAVVLDKLGQVLSQEQERYVLAENTRPSQMRKLFSLSQSMDRC 1409
 Qy 1397 KQGLYQALKETHPHILMELMEKSKKGLIPLSS 1429
 Db 1410 KQGLYQALKETHPHILMELMEKSKKGLIPLSS 1442
 RESULT 14
 ID AAY72673 standard; Protein; 1454 AA.
 AC AAY72673;
 DT 31-MAY-2001 (first entry)
 XX
 XX Human NAC beta isoform-CARD-X1 chimeric protein.
 DE Human: NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
 KW chimeric protein.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..1078
 FT /note="Corresponds to 1-1078 residues of human NAC
 FT beta isoform (AAY72669)"
 FT 1079..1454
 FT /note="Corresponds to 56-431 residues of human CARD-X
 FT protein (AAY72672)"
 FT
 PN MO200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000MO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Reed JC;
 PI
 XX WPI; 2001-183258/18.
 DR N-PSDB; AAD02764.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 PT
 PS Disclosure; Page 168-172; 184pp; English.
 XX
 CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
 CC beta isoform-CARD-X1 chimeric protein. NAC protein comprises a
 CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
 CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
 CC domain. CARD-X protein comprises a caspase-associated recruitment domain
 CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
 CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
 CC for screening modulators that modulates apoptosis, cytokine production,
 CC cytokine receptor signaling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory

CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukaemias.

SQ Sequence 1454 AA;

Query Match 84.2%; Score 6344; DB 22; Length 1454;
 Best Local Similarity 84.2%; Pred. No. 0;
 Matches 1229; Conservative 53; Mismatches 122; Indels 56; Gaps 7;

QY 1 MAGGAWRLACYLEPLKELKEFOLLANKASHRSSSGETPAQPKTSGMEVASYLVNQ 60
 DB 1 MAGGAWRLACYLEPLKELKEFOLLANKASHRSSSGETPAQPKTSGMEVASYLVNQ 60
 QY 61 YGEORAWDLALTWEQMGRLCAQAOEGAGSPSTPYSPSEPHLCSPOQSTAVLAMP 120
 DB 61 YGEORAWDLALTWEQMGRLCAQAOEGAGSPSTPYSPSEPHLCSPOQSTAVLAMP 120
 QY 121 IHELPAGCTQGSERRVLRQPTDSGRWRWEISASLYOALPSSPDHESPSQESPNAPTST 180
 DB 121 IHELPAGCTQGSERRVLRQPTDSGRWRWEISASLYOALPSSPDHESPSQESPNAPTST 180
 QY 181 AVLGWSGSPPOSLAREDEAPGTQWPLDETSGIYYTEIRERERESEKGRPPMAAVGT 240
 DB 181 AVLGWSGSPPOSLAREDEAPGTQWPLDETSGIYYTEIRERERESEKGRPPMAAVGT 240
 QY 241 PPOAHTSLOPHHHPWESPVSRESLCSTWPKNEDFNOKFTQLLIQPHPSODPLVKRSW 300
 DB 241 PPOAHTSLOPHHHPWESPVSRESLCSTWPKNEDFNOKFTQLLIQPHPSODPLVKRSW 300
 QY 301 PDVVENNRHLIEIRLPFGGLDTPRIVILQGAAGIGKSTLARKVEXAMGSGQLYGR 360
 DB 301 PDVVENNRHLIEIRLPFGGLDTPRIVILQGAAGIGKSTLARKVEXAMGSGQLYGR 360
 QY 361 FOHVEFYSCRELAQSKVSLAELIGDGTATPAPIRQILSRPERLLFIIDGVDEPGVQL 420
 DB 361 FOHVEFYSCRELAQSKVSLAELIGDGTATPAPIRQILSRPERLLFIIDGVDEPGVQL 420
 QY 421 EBSSESLCHWSQOPADALIGSLGKTIIPBASFLTARTALQNLIPSLBOARWVTVG 480
 DB 421 EBSSESLCHWSQOPADALIGSLGKTIIPBASFLTARTALQNLIPSLBOARWVTVG 480
 QY 481 FSESSEKREYFRFTBERQAIRAFVKSXKELMALCLVPWVSWLCTCLMOQMKREKL 540
 DB 481 FSESSEKREYFRFTBERQAIRAFVKSXKELMALCLVPWVSWLCTCLMOQMKREKL 540
 QY 541 TLTSTKTTTLCIHYLAQALQAPLGPQLNDLCSLAEGIWOKKTLFSPDLRKHGIDGAI 600
 DB 541 TLTSTKTTTLCIHYLAQALQAPLGPQLNDLCSLAEGIWOKKTLFSPDLRKHGIDGAI 600
 QY 601 ISTFLMGILIOEHPILPVSFHLCTQEFFPAAVSYLEDEKGRKSNCTIIDLEKTLAY 660
 DB 601 ISTFLMGILIOEHPILPVSFHLCTQEFFPAAVSYLEDEKGRKSNCTIIDLEKTLAY 660
 QY 661 GIHGLFGASTTRFLGILSDEGEREMENFHCRLSGGRNLMQVPSLQILLOPHSHESIA 720
 DB 661 GIHGLFGASTTRFLGILSDEGEREMENFHCRLSGGRNLMQVPSLQILLOPHSHESIA 720
 QY 721 CLVETENKTFELQVMAHFEMGCVETDEMLVCTFCIKFSRHVKYQLQLEGRORHSTWS 780
 DB 721 CLVETENKTFELQVMAHFEMGCVETDEMLVCTFCIKFSRHVKYQLQLEGRORHSTWS 780
 QY 781 PTWVVLFRWVPTVDAYWQILFSVLKATRNKKEIDLSGNSISHSAVXSLCKTLRRPCLLE 840
 DB 781 PTWVVLFRWVPTVDAYWQILFSVLKATRNKKEIDLSGNSISHSAVXSLCKTLRRPCLLE 840
 QY 841 TLRLAGCGTARECKOLAFGLRANQTLTELDSLFWLTDGAGHLCOBLRQPSCKORLO 900
 DB 841 TLRLAGCGTARECKOLAFGLRANQTLTELDSLFWLTDGAGHLCOBLRQPSCKORLO 900
 QY 901 LVSCGTLSDCCODLASYLSASPSLKEIDLQNNLMDVGVALCEGLRHAPACKLIRLGDLQ 960
 DB 901 LVSCGTLSDCCODLASYLSASPSLKEIDLQNNLMDVGVALCEGLRHAPACKLIRLGDLQ 960

QY 961 TTLSDEMOELALBOEKROLLIFSRKRSVMTPTREGIDTGMNSSTSLKORIGSERA 1020
 DB 961 TTLSDEMOELALBOEKROLLIFSRKRSVMTPTREGIDTGMNSSTSLKORIGSERA 1020
 QY 1021 ASHVAQANMLKLDVSKIFPIAIESSPEVVPVLLCVPSASQDGLTKPLGTDDEFW 1080
 DB 1021 ASHVAQANMLKLDVSKIFPIAIESSPEVVPVLLCVPSASQDGLTKPLGTDDEFW 1080
 QY 1081 GPTGPVATEVVDKKNLYRVHPVAGSYRMPNTGLCFVNRREAVYIEFCWVDQFLG-EI 1139
 DB 1081 GPTGPVATEVVDKKNLYRVHPVAGSYRMPNTGLCFVNRREAVYIEFCWVDQFLG-EI 1139
 QY 1140 NPOHSMVAGPLLDIKABE-GAVEAVHLPHFALQGNVDTSLFQMAHKEEGMILEXA 1198
 DB 1140 NPOHSMVAGPLLDIKABE-GAVEAVHLPHFALQGNVDTSLFQMAHKEEGMILEXA 1198
 QY 1141 QHHEQMLVGGPLFDVTAEBEAVABIHLPHTISLQ-GEVDVSWFLVAHKNGMWLEHFA 1199
 DB 1141 QHHEQMLVGGPLFDVTAEBEAVABIHLPHTISLQ-GEVDVSWFLVAHKNGMWLEHFA 1199
 QY 1199 RVELHHIYLVNPSFSPPLGVLLKMINNALPFIPTSVLLYHRVHPDEVFHLYLIPSDCS 1258
 DB 1199 RVELHHIYLVNPSFSPPLGVLLKMINNALPFIPTSVLLYHRVHPDEVFHLYLIPSDCS 1258
 QY 1200 KVEPPYAVLESFSPSLMGLRLRIASGTRLSIPTSTNTLITYHPHEDIKFHLIYVPSDAL 1259
 DB 1200 KVEPPYAVLESFSPSLMGLRLRIASGTRLSIPTSTNTLITYHPHEDIKFHLIYVPSDAL 1259
 QY 1259 IR-----KELECYRSPGSDQLF 1276
 DB 1259 IR-----KELECYRSPGSDQLF 1276
 QY 1260 LTKALIDDEDRPHVRLQTSPPMEPLNFGSSYIVSNSANLKVMPRELKLSYRSPGIQHF 1319
 DB 1260 LTKALIDDEDRPHVRLQTSPPMEPLNFGSSYIVSNSANLKVMPRELKLSYRSPGIQHF 1319
 QY 1277 SEFYGHLSGSLRLQVOKKQDETLYWEALYKRGDLMPTTLIPPARIAVPSLDAPQLH 1336
 DB 1277 SEFYGHLSGSLRLQVOKKQDETLYWEALYKRGDLMPTTLIPPARIAVPSLDAPQLH 1336
 QY 1320 SKFYAGQMKEPQLTEITERHGTIVMDTEVKVDQLVAASAP-----PFGGA-- 1369
 DB 1320 SKFYAGQMKEPQLTEITERHGTIVMDTEVKVDQLVAASAP-----PFGGA-- 1369
 QY 1337 FVDQYREOLIAVTSVEVVDLGLH-GOVLQSOBYRYLAENTRPSQMRKLFSLQSDWRK 1395
 DB 1337 FVDQYREOLIAVTSVEVVDLGLH-GOVLQSOBYRYLAENTRPSQMRKLFSLQSDWRK 1395
 QY 1370 FYKENHRQLQARMGDLKGVLDLQNEVLTENEKELVDEKTRQSKNEALLSMVEKKGD 1429
 DB 1370 FYKENHRQLQARMGDLKGVLDLQNEVLTENEKELVDEKTRQSKNEALLSMVEKKGD 1429
 QY 1396 CKDGLYQALKETHPHILMEL 1415
 DB 1396 CKDGLYQALKETHPHILMEL 1415
 QY 1430 ALDVLFRSISBRDPYLVSYL 1449
 DB 1430 ALDVLFRSISBRDPYLVSYL 1449

RESULT 15
 AA72674
 ID AA72674 standard; Protein; 1424 AA.
 XX
 AC AA72674;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Human NAC gamma or delta isoform-CARD-X1 chimeric protein.
 XX
 KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
 KW chimeric protein.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1048
 FT /note= "Corresponds to 1..1048 of human NAC delta
 (AA72671) or gamma (AA72670) isoform"
 FT Region 1049..1424
 FT /note= "Corresponds to 56-431 of human CARD-X
 protein (AA72672)"
 FT protein (AA72672)"
 PN WC0200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000MC-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 PA (BURN-) BURNHAM INST.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:03:41 ; Search time 1728.98 Seconds

(without alignments)
11475.528 Million cell updates/sec

Title: US-09-996-617-1

Perfect score: 1 gccccagggcctgagaggtc.....aggataagaagttactctac 5444

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4669878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0 %
Maximum Match 100 %
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5444	100.0	5444	10	US-09-996-617-1
2	5444	100.0	5444	10	US-09-931-071-1
3	5444	100.0	5444	11	US-09-956-712-3
4	5284.2	97.1	5531	11	US-09-956-712-11
5	4882.4	89.7	5100	11	US-09-956-712-12
6	4287	78.7	4287	10	US-09-996-617-5
7	4287	78.7	4287	10	US-09-931-071-5
8	4148	76.2	4422	10	US-09-388-221-1
9	4100	75.3	4200	10	US-09-388-221-3
10	3958	72.7	4332	10	US-09-388-221-5
11	3400.4	62.5	4556	10	US-09-388-221-9
12	3210.4	59.0	4466	10	US-09-388-221-11
13	2572.6	47.3	2657	11	US-09-895-298-22
14	1710.8	31.4	96649	11	US-09-956-712-10
15	694.4	12.8	720	13	US-10-027-632-150437

16	694.4	12.8	720	14	US-10-027-632-150437	Sequence 150437,
17	564	10.4	564	13	US-10-029-386-2806	Sequence 2806, Ap
18	508.8	9.3	725	13	US-10-027-632-20320	Sequence 20320, A
19	508.8	9.3	725	14	US-10-027-632-20320	Sequence 20320, A
20	489.4	9.0	491	10	US-09-796-692-5036	Sequence 5036, Ap
21	489.4	9.0	491	12	US-10-057-4758-5036	Sequence 5036, Ap
22	489.4	9.0	491	12	US-10-154-8848-5036	Sequence 5036, Ap
23	489.4	9.0	491	15	US-10-040-862-5036	Sequence 5036, Ap
24	484.4	8.9	578	11	US-09-764-891-1745	Sequence 1745, Ap
25	482.8	8.9	487	11	US-10-060-036-3299	Sequence 3299, Ap
26	414.4	7.6	416	15	US-10-060-036-3299	Sequence 3299, Ap
27	399.8	7.3	3186	15	US-10-066-521-17	Sequence 17, Appl
28	398.2	7.3	3465	12	US-10-108-2604-718	Sequence 718, App
29	398	7.3	3108	11	US-09-965-621-23	Sequence 23, Appl
30	398	7.3	3108	12	US-10-407-866-23	Sequence 23, Appl
31	398	7.3	3218	12	US-10-407-866-67	Sequence 67, Appl
32	369.8	6.8	3102	13	US-10-132-967-6	Sequence 6, Appl
33	369.8	6.8	3102	14	US-10-127-516-6	Sequence 6, Appl
34	369.8	6.8	3102	14	US-10-027-629-6	Sequence 6, Appl
35	369.8	6.8	3857	13	US-10-132-967-4	Sequence 4, Appl
36	369.8	6.8	3857	14	US-10-127-516-4	Sequence 4, Appl
37	369.8	6.8	3857	14	US-10-027-629-4	Sequence 4, Appl
38	340.6	6.3	371	12	US-10-242-535A-20893	Sequence 20893, A
39	318.4	5.8	322	10	US-09-796-692-2680	Sequence 2680, Ap
40	318.4	5.8	322	12	US-10-057-4758-2680	Sequence 2680, Ap
41	318.4	5.8	322	12	US-10-154-864B-2680	Sequence 2680, Ap
42	318.4	5.8	322	15	US-10-040-862-2680	Sequence 2680, Ap
43	313.4	5.8	2691	15	US-10-066-521-21	Sequence 21, Appl
44	292.8	5.4	2524	11	US-09-965-621-58	Sequence 58, Appl
45	292.8	5.4	2524	12	US-10-407-866-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-09-996-617-1
Sequence 1, Application US/09996617
Patent No. US2002012819A1
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-34001
CURRENT APPLICATION NUMBER: US/09/996.617
PRIOR FILING DATE: 2001-11-27
PRIOR APPLICATION NUMBER: 09/931,071
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/428,252
PRIOR FILING DATE: 1999-10-27
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 5444
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (523)...(4809)
US-09-996-617-1

Query Match 100.0%; Score 5444; DB 10; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGGGCTGAGAGCTGTAAGAACTGGAGGACAGCCGGGGCTCCACT 60
DB 1 GCCCAGGGCTGAGAGCTGTAAGAACTGGAGGACAGCCGGGGCTCCACT 60
QY 61 GGGTTCTAAGAGCCATTCCTGCTGCGGCTCCTCCACCCACCTCTTCAAGCTT 120

Dh 61 GGGTTCTGAAGCCCATCTCCCTGCTCTGGGGCTCTCCACCCCACTCTTCTCAGCTT 120
Qy 121 GCAGCTCAAGGGTGTATCTCAGAGATCCAGAGACCCAGAGAGGAGAAATCTGAGAA 180
Dh 121 GCAGCTCAAGGGTGTATCTCAGAGATCCAGAGACCCAGAGAGGAGAAATCTGAGAA 180
Qy 181 CAGAACGTAGAGGTTGCCCCACACCCCATCTCCGTACACATCTCCCTCACCTCAC 240
Dh 181 CAGAACGTAGAGGTTGCCCCACACCCCATCTCCGTACACATCTCCCTCACCTCAC 240
Qy 241 CCTCCGCTGAGCCCTGAGACCCCATCCAGAGACCTCCATATCAGTCAAGTCTGTCAGT 300
Dh 241 CCTCCGCTGAGCCCTGAGACCCCATCCAGAGACCTCCATATCAGTCAAGTCTGTCAGT 300
Qy 301 GTCTTGAGAGCCCTCTGAGGCTCTCCCTCCCTGAGCTTTTCCATCCACTCCCTCTAT 360
Dh 301 GTCTTGAGAGCCCTCTGAGGCTCTCCCTCCCTGAGCTTTTCCATCCACTCCCTCTAT 360
Qy 361 CGGCGTCTATCTGTAGGTGCTGGGATTTATAAACTGGGTTCCGATCTGAATTAAGA 420
Dh 361 CGGCGTCTATCTGTAGGTGCTGGGATTTATAAACTGGGTTCCGATCTGAATTAAGA 420
Qy 421 GAGGTAAGAGCCAGAGCAAGAGACAGTCTTCTGCTGCTGATACCTCCACAC 480
Dh 421 GAGGTAAGAGCCAGAGCAAGAGACAGTCTTCTGCTGCTGATACCTCCACAC 480
Qy 481 CTGGGAACATCCCCCAGACACCTCTTAACTCCGGAGACAGATGGCTGGGAGCTGG 540
Dh 481 CTGGGAACATCCCCCAGACACCTCTTAACTCCGGAGACAGATGGCTGGGAGCTGG 540
Qy 541 GGGCGCTGGCTGTATCTTGAAGTTCCTGAAGAGAGAGAGCTGAAGAGTTCCAGTT 600
Dh 541 GGGCGCTGGCTGTATCTTGAAGTTCCTGAAGAGAGAGAGCTGAAGAGTTCCAGTT 600
Qy 601 CTGCTGCGCAATAAAGGCACTCCAGAGGCTCTCCGTGAGACACCCCGTCCAGAGAG 660
Dh 601 CTGCTGCGCAATAAAGGCACTCCAGAGGCTCTCCGTGAGACACCCCGTCCAGAGAG 660
Qy 661 AAGACGAGTGCATGAGAGTGGCTCGTACCTGTGCTCAGTATGGGAGAGACGGGGCC 720
Dh 661 AAGACGAGTGCATGAGAGTGGCTCGTACCTGTGCTCAGTATGGGAGAGACGGGGCC 720
Qy 721 TGGGACCTTACCTCTCCTACCTGAGAGAGATGGGAGTGAAGTCACTGTGCGCCAGCC 780
Dh 721 TGGGACCTTACCTCTCCTACCTGAGAGAGATGGGAGTGAAGTCACTGTGCGCCAGCC 780
Qy 781 CAGGAGGGGAGGAGCACTCCCTCATGCTCCCTACAGCCAGAGTGAACCCCACTGGGG 840
Dh 781 CAGGAGGGGAGGAGCACTCCCTCATGCTCCCTACAGCCAGAGTGAACCCCACTGGGG 840
Qy 841 TCTCCAGCCCAACCTCCTCAGCGAGTGTATGCTGATCTCATGATTTGCGGG 900
Dh 841 TCTCCAGCCCAACCTCCTCAGCGAGTGTATGCTGATCTCATGATTTGCGGG 900
Qy 901 GGGGTGACCCAGGAGCTCAGAGAGAGGTTTGAAGACAGCTGCTGACATCTGAGCGC 960
Dh 901 GGGGTGACCCAGGAGCTCAGAGAGAGGTTTGAAGACAGCTGCTGACATCTGAGCGC 960
Qy 961 CGCTGAGAGAAATCTGCTGCTCATCTCCTACAGCTCTTCCAGAGTCCCAAGCAT 1020
Dh 961 CGCTGAGAGAAATCTGCTGCTCATCTCCTACAGCTCTTCCAGAGTCCCAAGCAT 1020
Qy 1021 GAGTCTCAAGAGAGAGTCAACCAAGCCCAATCAGACAGAGTGTGGGAGAGTGG 1080
Dh 1021 GAGTCTCAAGAGAGAGTCAACCAAGCCCAATCAGACAGAGTGTGGGAGAGTGG 1080
Qy 1081 GAGTCTCAAGAGAGAGTCAACCAAGCCCAATCAGACAGAGTGTGGGAGAGTGG 1140
Dh 1081 GAGTCTCAAGAGAGAGTCAACCAAGCCCAATCAGACAGAGTGTGGGAGAGTGG 1140
Qy 1141 CAGTCTGAGAGAGTCAAGAGATTTACTACAGAGAAATCAGAGAAAGAGAGAGAGAA 1200
Dh 1141 CAGTCTGAGAGAGTCAAGAGATTTACTACAGAGAAATCAGAGAAAGAGAGAGAGAA 1200

Qy 1201 TCAGAGAAAGGAGGCCCCCATAGGAGCGGTGTAGAAAGCCGCCCAAGGCGCACACC 1260
Dh 1201 TCAGAGAAAGGAGGCCCCCATAGGAGCGGTGTAGAAAGCCGCCCAAGGCGCACACC 1260
Qy 1261 AGCTTACAGCCCCCAGACCAACCATAGGAGCTCTCTGTAGAGAGAGGCTCTTCCACA 1320
Dh 1261 AGCTTACAGCCCCCAGACCAACCATAGGAGCTCTCTGTAGAGAGAGGCTCTTCCACA 1320
Qy 1321 TGGCCCTGGGAAAAATGAGATTTTAAACAAAAATTCACAAGCTGTCACTTACAAAAGA 1380
Dh 1321 TGGCCCTGGGAAAAATGAGATTTTAAACAAAAATTCACAAGCTGTCACTTACAAAAGA 1380
Qy 1381 CCTCACCCAGAGAGCCAGATCCCTGATCAAGAGAGTGGCTGTATATGTGAGAGAG 1440
Dh 1381 CCTCACCCAGAGAGCCAGATCCCTGATCAAGAGAGTGGCTGTATATGTGAGAGAG 1440
Qy 1441 AATCGAGACATTTAATTGAGATCAGAGCTTATTTGGGCCAGGCTGTGATCCCAAGAA 1500
Dh 1441 AATCGAGACATTTAATTGAGATCAGAGCTTATTTGGGCCAGGCTGTGATCCCAAGAA 1500
Qy 1501 CCTCGATAGTCACTAGAGGGGGCTGCTGGAAATTTGGGAAATTCACACTGGCAGAGCAG 1560
Dh 1501 CCTCGATAGTCACTAGAGGGGGCTGCTGGAAATTTGGGAAATTCACACTGGCAGAGCAG 1560
Qy 1561 GTGAAGAAAGCTGGGGAGAGGCGAGCTGTATGGGAGACGCTTCCAGCATGTCTTCTAC 1620
Dh 1561 GTGAAGAAAGCTGGGGAGAGGCGAGCTGTATGGGAGACGCTTCCAGCATGTCTTCTAC 1620
Qy 1621 TTCACTGCAAGAGAGTGGCCCACTGATCAAGTGTGTAGTCTGCTGAGCTCATCGAAAA 1680
Dh 1621 TTCACTGCAAGAGAGTGGCCCACTGATCAAGTGTGTAGTCTGCTGAGCTCATCGAAAA 1680
Qy 1681 GATGGAGAGCCACTCGGCTCCCATTTAGACAGATCTGTAGGCGAGAGCGGCTGCTC 1740
Dh 1681 GATGGAGAGCCACTCGGCTCCCATTTAGACAGATCTGTAGGCGAGAGCGGCTGCTC 1740
Qy 1741 TTCACTCTGATGTGTATGATGAGCCAGAGTGGGTTTTCAGAGAGCCGAGTTCTAGCTC 1800
Dh 1741 TTCACTCTGATGTGTATGATGAGCCAGAGTGGGTTTTCAGAGAGCCGAGTTCTAGCTC 1800
Qy 1801 TGTCTGACATGAGAGCCAGAGCAGCGGGGAGATGACCTCTGGGAGTTTGTCTGGGAAAA 1860
Dh 1801 TGTCTGACATGAGAGCCAGAGCAGCGGGGAGATGACCTCTGGGAGTTTGTCTGGGAAAA 1860
Qy 1861 ACTATACCTCCGAGGAGCATCTTCTGATCAAGGCTCGGAGCAGCTGTGAGAACTTC 1920
Dh 1861 ACTATACCTCCGAGGAGCATCTTCTGATCAAGGCTCGGAGCAGCTGTGAGAACTTC 1920
Qy 1921 ATTCTTCTTTGGAGAGAGCAGGTTGGGTAGAGTCTGGGGTTCTCTGAGTCCAGAGAG 1980
Dh 1921 ATTCTTCTTTGGAGAGAGCAGGTTGGGTAGAGTCTGGGGTTCTCTGAGTCCAGAGAG 1980
Qy 1981 AAGGAATATTTCTACAGATATTTACAGATGAAGGAGCAATTAAGACTTTAGATTG 2040
Dh 1981 AAGGAATATTTCTACAGATATTTACAGATGAAGGAGCAATTAAGACTTTAGATTG 2040
Qy 2041 GTCAAAATCAAAAGAGCTGAGGCTGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Dh 2041 GTCAAAATCAAAAGAGCTGAGGCTGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 TGCATTTGCTGATGAGAGAGTGAAGCGGAGGAAAACTCACTGACTTCAAGAGC 2160
Dh 2101 TGCATTTGCTGATGAGAGAGTGAAGCGGAGGAAAACTCACTGACTTCAAGAGC 2160
Qy 2161 ACCAGAACCTCTGTGTATTAATTAATTTGCGCAGGCTCTCAAGCTGAGCATTTGGGAGC 2220
Dh 2161 ACCAGAACCTCTGTGTATTAATTAATTTGCGCAGGCTCTCAAGCTGAGCATTTGGGAGC 2220
Qy 2221 CAGCTCAGAGAGCTCTGCTCTGCTGCTGAGGGGATCTGGCAAAAAAGACCTTTTC 2280
Dh 2221 CAGCTCAGAGAGCTCTGCTCTGCTGCTGAGGGGATCTGGCAAAAAAGACCTTTTC 2280

Oy 2281 AGTCAGATGACCTCAGGAGCATGGTTAGATGGGGCATCTCCACCTTCTTAA 2340
 Db 2281 AGTCAGATGACCTCAGGAGCATGGTTAGATGGGGCATCTCCACCTTCTTAA 2340
 Oy 2341 ATGGGATTTCTTCAAGAGCACCCTCTGAGCTACAGCTTCACTTCACTCTTTC 2400
 Db 2341 ATGGGATTTCTTCAAGAGCACCCTCTGAGCTACAGCTTCACTTCACTCTTTC 2400
 Oy 2401 CAAAGTTCTTTGACGAATGCTTATGCTTGGAGATGAGAGGGAGAGTAACAT 2460
 Db 2401 CAAAGTTCTTTGACGAATGCTTATGCTTGGAGATGAGAGGGAGAGTAACAT 2460
 Oy 2461 TCTAATGATCATGATGATTTGAAAAAGCTAGAGAGATAGAAATCATGGGCTGTT 2520
 Db 2461 TCTAATGATCATGATGATTTGAAAAAGCTAGAGAGATAGAAATCATGGGCTGTT 2520
 Oy 2521 GGGGATCAACCAACAGCTTCTTATTTGGGCTGTTAAGTATGAGGGGAGAGAGATG 2580
 Db 2521 GGGGATCAACCAACAGCTTCTTATTTGGGCTGTTAAGTATGAGGGGAGAGAGATG 2580
 Oy 2581 GAGAAATCTTTCACTGCGGCTGTCTCAAGGGAGAACTGATGCAAGTGGTCCGTC 2640
 Db 2581 GAGAAATCTTTCACTGCGGCTGTCTCAAGGGAGAACTGATGCAAGTGGTCCGTC 2640
 Oy 2641 CTGAGCTGCTGCTGACGACACCTCTGAGATGCTCCCACTGCTGTAAGAGCTGG 2700
 Db 2641 CTGAGCTGCTGCTGACGACACCTCTGAGATGCTCCCACTGCTGTAAGAGCTGG 2700
 Oy 2701 AACAAACGTTCTTGACACAAGATGAGCCCAATTTGAGAAATGGGATGTTAGAA 2760
 Db 2701 AACAAACGTTCTTGACACAAGATGAGCCCAATTTGAGAAATGGGATGTTAGAA 2760
 Oy 2761 AACAGATGAGACTTTTGTGTGACCTTTCTGCAATTAATGAGCCGCAAGTAAAG 2820
 Db 2761 AACAGATGAGACTTTTGTGTGACCTTTCTGCAATTAATGAGCCGCAAGTAAAG 2820
 Oy 2821 CTTGAGCTGATTTGAGGGAGGACGACAGATCAATGAGGCCCAATGGTATGCTG 2880
 Db 2821 CTTGAGCTGATTTGAGGGAGGACGACAGATCAATGAGGCCCAATGGTATGCTG 2880
 Oy 2881 TTCAAGTGGGCTCCAGTCAAGATGCTTATTTGGAGATTTCTTCTCCGCTCTAAG 2940
 Db 2881 TTCAAGTGGGCTCCAGTCAAGATGCTTATTTGGAGATTTCTTCTCCGCTCTAAG 2940
 Oy 2941 ACCAGAACTGAGAGAGCTGAGCTTAAGTGAATCTGCTGAGCCACTCTGAGTAAG 3000
 Db 2941 ACCAGAACTGAGAGAGCTGAGCTTAAGTGAATCTGCTGAGCCACTCTGAGTAAG 3000
 Oy 3001 AGCTTTGTAAGACCTGAGACGCTGCTGCTGCTGAGAGACCTGCGGTTGGCTGG 3060
 Db 3001 AGCTTTGTAAGACCTGAGACGCTGCTGCTGCTGAGAGACCTGCGGTTGGCTGG 3060
 Oy 3061 TGTGGCTTCAAGCTGAGAGCTGAGAGCTTGGCTTTGGGCTGAGAGCAACAGACC 3120
 Db 3061 TGTGGCTTCAAGCTGAGAGCTGAGAGCTTGGCTTTGGGCTGAGAGCAACAGACC 3120
 Oy 3121 CTGACCGAGCTGAGACTGAGCTTCAATGTGCTCAGAGATGCTGAGGCCAACA 3180
 Db 3121 CTGACCGAGCTGAGACTGAGCTTCAATGTGCTCAGAGATGCTGAGGCCAACA 3180
 Oy 3181 CAGAGACTGAGACGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 3240
 Db 3181 CAGAGACTGAGACGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 3240
 Oy 3241 ACCTCTGACTGCTGACAGAGCTGCTGCTGCTTATGCTGAGAGCCAGCTGAGAG 3300
 Db 3241 ACCTCTGACTGCTGACAGAGCTGCTGCTGCTTATGCTGAGAGCCAGCTGAGAG 3300
 Oy 3301 CTAGAGCTGAGAGCAACCTGATGACCTTGGCTGAGAGCTGCTGAGAGAGGCTC 3360
 Db 3301 CTAGAGCTGAGAGCAACCTGATGACCTTGGCTGAGAGCTGCTGAGAGAGGCTC 3360
 Oy 3361 AGGATCTGCTGAGAACTATAGCTGAGAGCTGAGAGCAACCTGATGATGAG 3420

Db 3361 AGGATCTGCTGAGAACTATAGCTGAGAGCTGAGAGCAACCTGATGATGAG 3420
 Oy 3421 ATGAGCAGAACTGAGGCTCTGAGCAGAGAACTCAGCTGCTCATCTTACAGAGA 3480
 Db 3421 ATGAGCAGAACTGAGGCTCTGAGCAGAGAACTCAGCTGCTCATCTTACAGAGA 3480
 Oy 3481 CGGAAACCAATGTGATGACCTTATCTGAGGCTCTGATACGGAGAGATGATATG 3540
 Db 3481 CGGAAACCAATGTGATGACCTTATCTGAGGCTCTGATACGGAGAGATGATATG 3540
 Oy 3541 ACATCTCACTCAAGCCGACAGACTGAGATCAGAGAGGGGCTTCCATGTTGCTAG 3600
 Db 3541 ACATCTCACTCAAGCCGACAGACTGAGATCAGAGAGGGGCTTCCATGTTGCTAG 3600
 Oy 3601 GCTAATCTCAAACTCTGAGAGTGAAGCAATCTTCCATTTGCTGATGACAGAGAA 3660
 Db 3601 GCTAATCTCAAACTCTGAGAGTGAAGCAATCTTCCATTTGCTGATGACAGAGAA 3660
 Oy 3661 AGCTCCCAAGAGTATACCGGTGAATCTTGTGCGTCTTCTGCTCTCAAGGG 3720
 Db 3661 AGCTCCCAAGAGTATACCGGTGAATCTTGTGCGTCTTCTGCTCTCAAGGG 3720
 Oy 3721 GACCTGCAATAGAAAGCTTTGGGAGCTGAGATGATCTTGGGGGCCACAGGGGCTG 3780
 Db 3721 GACCTGCAATAGAAAGCTTTGGGAGCTGAGATGATCTTGGGGGCCACAGGGGCTG 3780
 Oy 3781 GCTACTGAGTATGACAAAGAAAGAACTTGTACCGAGTTCACTTCTGATGCTGG 3840
 Db 3781 GCTACTGAGTATGACAAAGAAAGAACTTGTACCGAGTTCACTTCTGATGCTGG 3840
 Oy 3841 TCTTACCGCTGGCCCAACAGGGTCTGCTTTGTGATGAGAGAGCGGTGAG 3900
 Db 3841 TCTTACCGCTGGCCCAACAGGGTCTGCTTTGTGATGAGAGAGCGGTGAG 3900
 Oy 3901 ATTGAATCTGTGTGTGGAGCAAGTCTGAGTGAATCAACCAAGAGCTGATG 3960
 Db 3901 ATTGAATCTGTGTGTGGAGCAAGTCTGAGTGAATCAACCAAGAGCTGATG 3960
 Oy 3961 GTGGAGGGCTCTGCTGAGCATCAAGCTGAGCTGAGAGCTGAGAGCTGAGAGCT 4020
 Db 3961 GTGGAGGGCTCTGCTGAGCATCAAGCTGAGCTGAGAGCTGAGAGCTGAGAGCT 4020
 Oy 4021 CCTCACTTTGGCTCTTCAAGGGGCTGATGAGACATCTCCCTGTTCAATGGCCAC 4080
 Db 4021 CCTCACTTTGGCTCTTCAAGGGGCTGATGAGACATCTCCCTGTTCAATGGCCAC 4080
 Oy 4081 TTTAAGAGAGGGAGATGCTCTGAGAGAGCAGCAGGGGAGAGCTGATCACTAGT 4140
 Db 4081 TTTAAGAGAGGGAGATGCTCTGAGAGAGCAGCAGGGGAGAGCTGATCACTAGT 4140
 Oy 4141 CTGAGAAACCCAGCTTCTCCCTTGGAGTCTCTGAAATGATCATATGCTGCTG 4200
 Db 4141 CTGAGAAACCCAGCTTCTCCCTTGGAGTCTCTGAAATGATCATATGCTGCTG 4200
 Oy 4201 CGCTTCAATCCCGTCACTCTGTGTGTGCTTTTACACCGGCTTCACTGAGAGAT 4260
 Db 4201 CGCTTCAATCCCGTCACTCTGTGTGTGCTTTTACACCGGCTTCACTGAGAGAT 4260
 Oy 4261 ACCTTCAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320
 Db 4261 ACCTTCAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320
 Oy 4321 TATGAGAGCTTGAAGAGACAGCTGCTTGAAGTCTTCACTTGGCTGATGATGAT 4380
 Db 4321 TATGAGAGCTTGAAGAGACAGCTGCTTGAAGTCTTCACTTGGCTGATGATGAT 4380
 Oy 4381 GGGATCAGGCTGAG 4440
 Db 4381 GGGATCAGGCTGAG 4440
 Oy 4441 AAACAGAGATCTCATGCTGCAACTACTGATCTTCCAGGCTGAGAGAGAGAGAG 4500

D 721 TGGAGCTTACGCTCCATACCTGGAGACAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC 780
Q 781 CAGGAAGGGCAGGCGCACTCTCCCTCAATCCCTTACAGCCCAAGTGAACCCCACTTGGG 840
D 781 CAGGAAGGGCAGGCGCACTCTCCCTCAATCCCTTACAGCCCAAGTGAACCCCACTTGGG 840
Q 841 TCTCCCAAGCCCACTCTCCCTCAATCCCTTACAGCCCAAGTGAACCCCACTTGGG 900
D 841 TCTCCCAAGCCCACTCTCCCTCAATCCCTTACAGCCCAAGTGAACCCCACTTGGG 900
Q 901 GGGGACCCAGGGCTCAGAGAGAGAGGTTTGAAGAGGCTGTGACATCTGTGAGCG 960
D 901 GGGGACCCAGGGCTCAGAGAGAGAGGTTTGAAGAGGCTGTGACATCTGTGAGCG 960
Q 961 CGCTGAGAGAGAAATCTCTGCTCACTCTTACCAAGCTCTTCCAGCTCCCAAGCAT 1020
D 961 CGCTGAGAGAGAAATCTCTGCTCACTCTTACCAAGCTCTTCCAGCTCCCAAGCAT 1020
Q 1021 GAGTCTTCAAGCAGAGAGTCAACCAAGCCCAATCCACAGCATGTGAGGAGCTGG 1080
D 1021 GAGTCTTCAAGCAGAGAGTCAACCAAGCCCAATCCACAGCATGTGAGGAGCTGG 1080
Q 1081 GGAATCCCACTCAGCCAGCTAGCAACCAAGAGAGAGAGGCTCTGGGAGCCAAATG 1140
D 1081 GGAATCCCACTCAGCCAGCTAGCAACCAAGAGAGAGAGGCTCTGGGAGCCAAATG 1140
Q 1141 CCTCTGATGAAAGCTCAGAGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAA 1200
D 1141 CCTCTGATGAAAGCTCAGAGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAA 1200
Q 1201 TCGAGAGAGAGAGAGGCCCCATGGGAGCGGTGTGAGAAAGCCCCCAAGGCGCAACC 1260
D 1201 TCGAGAGAGAGAGAGGCCCCATGGGAGCGGTGTGAGAAAGCCCCCAAGGCGCAACC 1260
Q 1261 AGGCTCAGGCCCCAAGCAACCAAGAGAGGCTTGTGAGAGAGAGGCTGTGAGAG 1320
D 1261 AGGCTCAGGCCCCAAGCAACCAAGAGAGGCTTGTGAGAGAGAGGCTGTGAGAG 1320
Q 1321 TGGCCCTGAGAAATGAGAGATTTTAAACAAAAATTCACAGAGCTGCTACTTCAAGA 1380
D 1321 TGGCCCTGAGAAATGAGAGATTTTAAACAAAAATTCACAGAGCTGCTACTTCAAGA 1380
Q 1381 CCTCAGCCCAAGAGCCCAAGATCCCTGTGTCAAGAGAGCTGCTGTATGTGAGAG 1440
D 1381 CCTCAGCCCAAGAGCCCAAGATCCCTGTGTCAAGAGAGCTGCTGTATGTGAGAG 1440
Q 1441 AATCAGAGACATTTAATTGAGATCAGAGACTTAATTTGGCCCAAGGCTGATACCAAGA 1500
D 1441 AATCAGAGACATTTAATTGAGATCAGAGACTTAATTTGGCCCAAGGCTGATACCAAGA 1500
Q 1501 CCTCAGATGATCATCTGCAAGGGGCTGTGAAATTTGGAAATCACTGCGCCAGAG 1560
D 1501 CCTCAGATGATCATCTGCAAGGGGCTGTGAAATTTGGAAATCACTGCGCCAGAG 1560
Q 1561 GTGAAGAGAGCTGGGGGAGAGGCTGATGTGAGGAGCGCTTCCAGCATGTCTTAC 1620
D 1561 GTGAAGAGAGCTGGGGGAGAGGCTGATGTGAGGAGCGCTTCCAGCATGTCTTAC 1620
Q 1621 TTTAGCTGAGAGAGCTGGGCGCAGTCCAAGGTGTGAGTCTCGTGAAGCTATCGGAAA 1680
D 1621 TTTAGCTGAGAGAGCTGGGCGCAGTCCAAGGTGTGAGTCTCGTGAAGCTATCGGAAA 1680
Q 1681 GATGAGACAGCACTCGGCTCCCATTTAGACAGATCTGTCTAGGCGCAGAGCGGCTCTC 1740
D 1681 GATGAGACAGCACTCGGCTCCCATTTAGACAGATCTGTCTAGGCGCAGAGCGGCTCTC 1740
Q 1741 TTTATCTCTGATGTGTGATGATGAGCCAGATGGGTCTTTCAGAGAGCGAGTTCTGAGCTC 1800
D 1741 TTTATCTCTGATGTGTGATGATGAGCCAGATGGGTCTTTCAGAGAGCGAGTTCTGAGCTC 1800
Q 1801 TGTCTGAGCTGAGAGCGCAAGCGCGGAGTGAATGCTGTGGCAAGTTTGTGGGAGAA 1860
D 1801 TGTCTGAGCTGAGAGCGCAAGCGCGGAGTGAATGCTGTGGCAAGTTTGTGGGAGAA 1860

Q 1861 ACTATATCTTCCAGGAGCATCTTCTGATCAGGCTCGAGCAAGCTCTGCAAGACCTC 1920
D 1861 ACTATATCTTCCAGGAGCATCTTCTGATCAGGCTCGAGCAAGCTCTGCAAGACCTC 1920
Q 1921 ATTCTCTTCTTGAAGAGAGAGCTGGGTGAGAGTCTTGGGGTCTTCTGATCTCAGAGG 1980
D 1921 ATTCTCTTCTTGAAGAGAGAGCTGGGTGAGAGTCTTGGGGTCTTCTGATCTCAGAGG 1980
Q 1981 AAGGAATATTTTACAGATATTTTACAGAGTGAAGGCAAGCAATTTAGAGCTTTAGGTTG 2040
D 1981 AAGGAATATTTTACAGATATTTTACAGAGTGAAGGCAAGCAATTTAGAGCTTTAGGTTG 2040
Q 2041 GTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCTGTGGGTCTGTGGCTGGCC 2100
D 2041 GTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCTGTGGGTCTGTGGCTGGCC 2100
Q 2101 TGCATCTTGGCTGATGAGAGAGATGAAGCGGAGAGAAATCTCACATGATTTCCAGAGC 2160
D 2101 TGCATCTTGGCTGATGAGAGAGATGAAGCGGAGAGAAATCTCACATGATTTCCAGAGC 2160
Q 2161 ACCACAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAGCTCAGCCATTGGAGCC 2220
D 2161 ACCACAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAGCTCAGCCATTGGAGCC 2220
Q 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGGGATCTGGCAAAAAGAACCTTTTC 2280
D 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGGGATCTGGCAAAAAGAACCTTTTC 2280
Q 2281 AGTCCAGATGACCTCAGAGAGAGATGGGTGATGAGGGGCTCATCTTCACTTCTTGAAG 2340
D 2281 AGTCCAGATGACCTCAGAGAGAGATGGGTGATGAGGGGCTCATCTTCACTTCTTGAAG 2340
Q 2341 ATGGATATTTTCAAGAGACCCCATCCCTGTAGCTACAGCTTCACTCTCTGTTTC 2400
D 2341 ATGGATATTTTCAAGAGACCCCATCCCTGTAGCTACAGCTTCACTCTCTGTTTC 2400
Q 2401 CAAAGGTTCTTTTGCAGCAATGTCTTATGTGAGAGATGAGAGGGAGAGGTTAAACAT 2460
D 2401 CAAAGGTTCTTTTGCAGCAATGTCTTATGTGAGAGATGAGAGGGAGAGGTTAAACAT 2460
Q 2461 TCTAATTTGATCATTAATTTGAGAAAGAGCTGTGAAGCAATGAGAAATCAATGCGCTGTT 2520
D 2461 TCTAATTTGATCATTAATTTGAGAAAGAGCTGTGAAGCAATGAGAAATCAATGCGCTGTT 2520
Q 2521 GGGGATCAACCAAGCTTCTTATTTGGGCTGTAAAGTATGAGGGGAGAGAGATG 2580
D 2521 GGGGATCAACCAAGCTTCTTATTTGGGCTGTAAAGTATGAGGGGAGAGAGATG 2580
Q 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGAGAGAACCTGATGCAATGGTCCGCTC 2640
D 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGAGAGAACCTGATGCAATGGTCCGCTC 2640
Q 2641 CTGCACTGTGCTGAGAGCCCACTCTGTGAATGCTTCCAGCTGTTTGAAGAGCTCGG 2700
D 2641 CTGCACTGTGCTGAGAGCCCACTCTGTGAATGCTTCCAGCTGTTTGAAGAGCTCGG 2700
Q 2701 AACAAAGCTTCTGACACAGATGAGCCATTTTGAAGAAATGGGATGTGTGAA 2760
D 2701 AACAAAGCTTCTGACACAGATGAGCCATTTTGAAGAAATGGGATGTGTGAA 2760
Q 2761 ACAGACATGAGAGCTTTAGTGTGCACTTTCTGATTAATTTAGCGCCACAGTGAAGAG 2820
D 2761 ACAGACATGAGAGCTTTAGTGTGCACTTTCTGATTAATTTAGCGCCACAGTGAAGAG 2820
Q 2821 CTTCACTGATTTAAGGAGAGAGAGCAAGATCAATGAGAGAGAGAGAGAGAGAGAGAG 2880
D 2821 CTTCACTGATTTAAGGAGAGAGAGCAAGATCAATGAGAGAGAGAGAGAGAGAGAGAG 2880
Q 2881 TTTAGGAGGAGCTCAGATGAGATGCTATTTGAGAGATTTCTTCTCCGCTCTCAAGCTC 2940
D 2881 TTTAGGAGGAGCTCAGATGAGATGCTATTTGAGAGATTTCTTCTCCGCTCTCAAGCTC 2940

QY 2941 ACCAGAAACCTGAGAGAGCTGACCTTAAGTGAACCTCGCTGAGCCACTCTGACGTGAAG 3000
 DB 2941 ACCAGAAACCTGAGAGAGCTGACCTTAAGTGAACCTCGCTGAGCCACTCTGAGTGAAG 3000
 QY 3001 AGCTCTTGTAAAGCCTGAGAGAGCCTCGCTGCTCTGAGAGAGCCTGCGGTTGCTGAGC 3060
 DB 3001 AGCTCTTGTAAAGCCTGAGAGAGCCTCGCTGCTCTGAGAGAGCCTGCGGTTGCTGAGC 3060
 QY 3061 TGTGAGCTCAAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGAGAGC 3120
 DB 3061 TGTGAGCTCAAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGAGAGC 3120
 QY 3121 CTGAGCCGAGCTGAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGAGC 3180
 DB 3121 CTGAGCCGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGC 3180
 QY 3181 CAGAGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGAGC 3240
 DB 3181 CAGAGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGAGC 3240
 QY 3241 AGCTCTGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGAGC 3300
 DB 3241 AGCTCTGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTTGTGCTGAGAGAGAGAGC 3300
 QY 3301 CTGAGAGCTGAGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGC 3360
 DB 3301 CTGAGAGCTGAGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGC 3360
 QY 3361 AGGAGCTGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGC 3420
 DB 3361 AGGAGCTGAGCTGAGAGAGCTGAGAGAGCTGCAAGAGAGCTGCAAGAGAGCTGCAAGAGC 3420
 QY 3421 ATGAGAGAGAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGAGC 3480
 DB 3421 ATGAGAGAGAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGAGC 3480
 QY 3481 CGGAGAGAGAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGAGC 3540
 DB 3481 CGGAGAGAGAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGAGC 3540
 QY 3541 ACATCTCTCACTCAAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGC 3600
 DB 3541 ACATCTCTCACTCAAGAGAGAGCTGAGAGAGCTGCAAGAGAGAGCTTGTGCTGAGAGAGC 3600
 QY 3601 GCTAATCTCACTCAAGAGAGAGCTGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGC 3660
 DB 3601 GCTAATCTCACTCAAGAGAGAGCTGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGC 3660
 QY 3661 AGCTCTCTCACTCAAGAGAGAGCTGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGC 3720
 DB 3661 AGCTCTCTCACTCAAGAGAGAGCTGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGC 3720
 QY 3721 GACCTCTCACTCAAGAGAGAGCTGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGC 3780
 DB 3721 GACCTCTCACTCAAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTG 3780
 QY 3781 GCTAATCTCACTCAAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGC 3840
 DB 3781 GCTAATCTCACTCAAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGC 3840
 QY 3841 TCTTACCGCTGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGC 3900
 DB 3841 TCTTACCGCTGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGC 3900
 QY 3901 ATTGAATCTGAGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGC 3960
 DB 3901 ATTGAATCTGAGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGC 3960
 QY 3961 GTGAGAGAGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGAGC 4020
 DB 3961 GTGAGAGAGAGAGAGAGCTTGTGCTGAGAGAGAGCTTGTGCTGAGAGAGAGAGAGC 4020
 QY 4021 CCTCACTTGTGAGC 4080

DB 4021 CCTCACTTGTGAGC 4080
 QY 4081 TTTAAAG 4140
 DB 4081 TTTAAAG 4140
 QY 4141 CTGAG 4200
 DB 4141 CTGAG 4200
 QY 4201 GAGCTTCACTTGTGAGC 4260
 DB 4201 GAGCTTCACTTGTGAGC 4260
 QY 4261 ACCCTTCACTTGTGAGC 4320
 DB 4261 ACCCTTCACTTGTGAGC 4320
 QY 4321 TATCGAG 4380
 DB 4321 TATCGAG 4380
 QY 4381 GAGATCAG 4440
 DB 4381 GAGATCAG 4440
 QY 4441 AAACAG 4500
 DB 4441 AAACAG 4500
 QY 4501 TCACTTGTGAG 4560
 DB 4501 TCACTTGTGAG 4560
 QY 4561 GCCCGAG 4620
 DB 4561 GCCCGAG 4620
 QY 4621 GAG 4680
 DB 4621 GAG 4680
 QY 4681 AGCTTGAAG 4740
 DB 4681 AGCTTGAAG 4740
 QY 4741 ACCCATCTCACTTGTGAG 4800
 DB 4741 ACCCATCTCACTTGTGAG 4800
 QY 4801 CTGAG 4860
 DB 4801 CTGAG 4860
 QY 4861 TTTCTTGGGTCTGAG 4920
 DB 4861 TTTCTTGGGTCTGAG 4920
 QY 4921 CTAAAGTAATGAG 4980
 DB 4921 CTAAAGTAATGAG 4980
 QY 4981 CCAAG 5040
 DB 4981 CCAAG 5040
 QY 5041 GGAAG 5100
 DB 5041 GGAAG 5100
 QY 5101 GCCCTTCCGAG 5160

Db 5101 GCCCTCGATCCCATCCAGCGGCAAGAGAAATAGAGGGAATGGAACCATTTGCTC 5160
QY 5161 TGGCTGTGTACAGGGGTGAGCCCCCAAAATTTGGGTTTACGCTGGAGGCCAGTGTATTC 5220
Db 5161 TGGCTGTGTACAGGGGTGAGCCCCCAAAATTTGGGTTTACGCTGGAGGCCAGTGTATTC 5220
QY 5221 TTGGCTTTGTACAGGAAGATCTACAGAGAGCCACAGAGTAAGTGAAGAAAGTTT 5280
Db 5221 TTGGCTTTGTACAGGAAGATCTACAGAGAGCCACAGAGTAAGTGAAGAAAGTTT 5280
QY 5281 ATTCAAGAAATTAAGAGATATCAAGCTCTTTTGAATTTGTCTAGAGGCTTTCCAGTT 5340
Db 5281 ATTCAAGAAATTAAGAGATATCAAGCTCTTTTGAATTTGTCTAGAGGCTTTCCAGTT 5340
QY 5341 TTATACAGAAACCCCTATATAATTTAAATTTTAACTTAATTAAGAAATTAAGAAAT 5400
Db 5341 TTATACAGAAACCCCTATATAATTTAAATTTTAACTTAATTAAGAAATTAAGAAAT 5400
QY 5401 ACAAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444
Db 5401 ACAAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444

RESULT 3
US-09-956-712-3

/ Sequence 3, Application US/09956712
/ Publication No. US2003092648A1
/ GENERAL INFORMATION:
/ APPLICANT: C. Frank Bennett
/ APPLICANT: Susan M. Freiler
/ TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
/ FILE REFERENCE: RTS-0326
/ CURRENT APPLICATION NUMBER: US/09/956,712
/ NUMBER OF SEQ ID NOS: 91
/ SEQ ID NO 3
/ LENGTH: 5444
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (523)...(4812)
US-09-956-712-3

Query Match 100.0%; Score 5444; DB 11; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCCAGGGCTGGAAGATCTGAAGAAACCTGGAGAGCCAGAGCCCGGGGCTTCACTCT 60
Db 1 GCCCCAGGGCTGGAAGATCTGAAGAAACCTGGAGAGCCAGAGCCCGGGGCTTCACTCT 60
QY 61 GGGTTCTGAAGAGCCATCTCTGCTCTGCGGCTCTTCCACCCCACTTCTCAGGCTT 120
Db 61 GGGTTCTGAAGAGCCATCTCTGCTCTGCGGCTCTTCCACCCCACTTCTCAGGCTT 120
QY 121 GCAGCTCAAGGGTTGATCTCAGAGATCAGAGAGCCAGAGAGAGGAGAAATCTGAGAA 180
Db 121 GCAGCTCAAGGGTTGATCTCAGAGATCAGAGAGCCAGAGAGAGGAGAAATCTGAGAA 180
QY 181 CAGAAAGTGAAGGTGAGCCAGAGAGCCATCTCCGCTACAGATCTCCCTCCAGCCCTAC 240
Db 181 CAGAAAGTGAAGGTGAGCCAGAGAGCCATCTCCGCTACAGATCTCCCTCCAGCCCTAC 240
QY 241 CCTCCCTGAGGAGCCCTGAGAGCCATCTCCAGAGAGCCCTCTATCAGTGTCTTCCAGT 300
Db 241 CCTCCCTGAGGAGCCCTGAGAGCCATCTCCAGAGAGCCCTCTATCAGTGTCTTCCAGT 300
QY 301 GTCTTGAAGAGCCCTGAGGAGCCCTCTCCCTCCCTGAGGAGGAGGAGGAGGAGGAG 360
Db 301 GTCTTGAAGAGCCCTGAGGAGCCCTCTCCCTCCCTGAGGAGGAGGAGGAGGAGGAG 360
QY 361 CGAGGCTATCTGTAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

Db 361 CGAGGCTATCTGTAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 GAGGTGAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 421 GAGGTGAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 CTGGAGAAATCATCCCCAGAGAGCCCTTTAACTCCGAGAGAGAGATGCTGGAGAGCTTG 540
Db 481 CTGGAGAAATCATCCCCAGAGAGCCCTTTAACTCCGAGAGAGAGATGCTGGAGAGCTTG 540
QY 541 GGGGCTGGGCTGTTAGTCTTGGAGTCTGAGAGAGAGAGGAGGAGGAGGAGGAGGAG 600
Db 541 GGGGCTGGGCTGTTAGTCTTGGAGTCTGAGAGAGAGAGGAGGAGGAGGAGGAGGAG 600
QY 601 CTGCTGCGCAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 601 CTGCTGCGCAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 AAGACAGATGAGCATGAGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 661 AAGACAGATGAGCATGAGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 TGGGAGCTATGAGGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 721 TGGGAGCTATGAGGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 TCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 841 TCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 GGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 901 GGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 CGCTGAGAGAAATCTCTGCTCTCACTCTTCAAGAGCTTTTCAAGCTTCCAGAGCAT 1020
Db 961 CGCTGAGAGAAATCTCTGCTCTCACTCTTCAAGAGCTTTTCAAGCTTCCAGAGCAT 1020
QY 1021 GAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1021 GAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 GAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 CCTCTGATGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 CCTCTGATGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 AGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 AGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 TGGGCTGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 TGGGCTGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 1381 CTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381 CTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 AATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

D 1441 AATCGAGCATTTAATGAGATCAAGACTTAATTTGGCCAGGCTTGATACCAAGAA 1500
Q 1501 CCTGCAATGATCACTGAGGAGGAGCTGCTGGAATTGGGAATGCAACTGAGCAG 1560
D 1501 CCTGCAATGATCACTGAGGAGGAGCTGCTGGAATTGGGAATGCAACTGAGCAG 1560
Q 1561 GTGAAGAAAGCTGGGGGAGAGGAGCTGATAGGGAGCCGCTTCCAGAGTCTTCTAC 1620
D 1561 GTGAAGAAAGCTGGGGGAGAGGAGCTGATAGGGAGCCGCTTCCAGAGTCTTCTAC 1620
Q 1621 TTGAGCTGAGAGAGCTGAGCCCAAGTGGTGTGAGTCTGCTGAGCTCATCGGAAA 1680
D 1621 TTGAGCTGAGAGAGCTGAGCCCAAGTGGTGTGAGTCTGCTGAGCTCATCGGAAA 1680
Q 1681 GATGGAGAGCACTCCGAGCTCCCAATTAAGACAGATCTGTCTAGGCGAGGAGTCTC 1740
D 1681 GATGGAGAGCACTCCGAGCTCCCAATTAAGACAGATCTGTCTAGGCGAGGAGTCTC 1740
Q 1741 TTGATCTGATGATGATGAGTGGTGTGAGTGGTGTGAGTGGTGTGAGTGGTGTGAG 1800
D 1741 TTGATCTGATGATGATGAGTGGTGTGAGTGGTGTGAGTGGTGTGAGTGGTGTGAG 1800
Q 1801 TGTCTGATGAGAGCAAGCAAGCCGAGATGCACTGCTGGGAGTGTGCTGGGAAA 1860
D 1801 TGTCTGATGAGAGCAAGCAAGCCGAGATGCACTGCTGGGAGTGTGCTGGGAAA 1860
Q 1861 ACTAATCTCCGAGGAGCTCTCTGATCAAGCTCTGAGCAAGCTCTGAGAACTTC 1920
D 1861 ACTAATCTCCGAGGAGCTCTCTGATCAAGCTCTGAGCAAGCTCTGAGAACTTC 1920
Q 1921 ATTCTCTTTGAGAGGAGGAGCTGGTGTGAGTGGTGTGAGTGGTGTGAGTGGTGTGAG 1980
D 1921 ATTCTCTTTGAGAGGAGGAGCTGGTGTGAGTGGTGTGAGTGGTGTGAGTGGTGTGAG 1980
Q 1981 AAGGAATATTTCTACAGATATTTCTACAGATGAAAGCAAGCAATTAGAGCTTGA 2040
D 1981 AAGGAATATTTCTACAGATATTTCTACAGATGAAAGCAAGCAATTAGAGCTTGA 2040
Q 2041 GTCAATCAAAAGAGCTCTGAGGCTGTGTCTGTGCTGTGCTGTGCTGTGCTGTGCT 2100
D 2041 GTCAATCAAAAGAGCTCTGAGGCTGTGTCTGTGCTGTGCTGTGCTGTGCTGTGCT 2100
Q 2101 TGCACTGCTGAGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
D 2101 TGCACTGCTGAGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Q 2161 ACCAAGACCTCTGTCTACATTAATCTGAGGCTCTGAGGCTCTGAGGCTCTGAGG 2220
D 2161 ACCAAGACCTCTGTCTACATTAATCTGAGGCTCTGAGGCTCTGAGGCTCTGAGG 2220
Q 2221 CAGCTGAGAGGCTCTGTCTCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
D 2221 CAGCTGAGAGGCTCTGTCTCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Q 2281 AGTCAATGAGCTCAGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
D 2281 AGTCAATGAGCTCAGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Q 2341 ATGGGATATTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
D 2341 ATGGGATATTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
Q 2401 CAGAGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
D 2401 CAGAGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
Q 2461 TCTAATGAGCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
D 2461 TCTAATGAGCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
Q 2521 GGGGATCAACCAAGCTTTCTATTTGGGCTGTGTAAGTGAAGGAGGAGGAGGAGG 2580
D 2521 GGGGATCAACCAAGCTTTCTATTTGGGCTGTGTAAGTGAAGGAGGAGGAGGAGG 2580

Q 2581 GAGAACTTTTCACTGCGGCTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
D 2581 GAGAACTTTTCACTGCGGCTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
Q 2641 CTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700
D 2641 CTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700
Q 2701 AACAAAAGCTTCTGAGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
D 2701 AACAAAAGCTTCTGAGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
Q 2761 ACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
D 2761 ACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
Q 2821 CTTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2880
D 2821 CTTGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2880
Q 2881 TTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
D 2881 TTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
Q 2941 ACCAAGAACTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3000
D 2941 ACCAAGAACTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3000
Q 3001 AGTCTTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060
D 3001 AGTCTTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060
Q 3061 TGTGAGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3120
D 3061 TGTGAGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3120
Q 3121 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180
D 3121 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180
Q 3181 CAGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
D 3181 CAGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
Q 3241 AGCTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3300
D 3241 AGCTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3300
Q 3301 CTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360
D 3301 CTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360
Q 3361 AGGATCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420
D 3361 AGGATCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420
Q 3421 ATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
D 3421 ATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
Q 3481 CAGAAACCAAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
D 3481 CAGAAACCAAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
Q 3541 ACATCTCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600
D 3541 ACATCTCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600
Q 3601 GCTAATCTCAACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660
D 3601 GCTAATCTCAACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660

QY	3661	AGCTACCCAGAGAGTACCGGTGGAACTCTTGCGCGGCTTCCTCCGCTCCAAAGG	3720
Db	3661	AGCTACCCAGAGAGTACCGGTGGAACTCTTGCGCGGCTTCCTCCGCTCCAAAGG	3720
QY	3721	GACCTGCAATCGAAGCCTTTGGGGACATGACGANTGACTTTGGGGGCCCCACCGGGCCCTGTG	3780
Db	3721	GACCTGCAATCGAAGCCTTTGGGGACATGACGANTGACTTTGGGGGCCCCACCGGGCCCTGTG	3780
QY	3781	GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC	3840
Db	3781	GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC	3840
QY	3841	TCCATACGGCTGGGCCCAACAGGGGTCTGCTCTTGATGTAAGAAAGCGGTACCGTTGAG	3900
Db	3841	TCCATACGGCTGGGCCCAACAGGGGTCTGCTCTTGATGTAAGAAAGCGGTACCGTTGAG	3900
QY	3901	ATTGAATTCGTGTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGCAGCTGGATG	3960
Db	3901	ATTGAATTCGTGTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGCAGCTGGATG	3960
QY	3961	GTGGCAGGGGCTGTGCTGGACATCAAGCTGAGCCTGGAGCTGTGGAAGCTGTGCACTTC	4020
Db	3961	GTGGCAGGGGCTGTGCTGGACATCAAGCTGAGCCTGGAGCTGTGGAAGCTGTGCACTTC	4020
QY	4021	CCTCACTTTGTGTGTCTTCCAAAGGGGGCCATGTGGACACATCTCCCTGTCCAAATGGCCAC	4080
Db	4021	CCTCACTTTGTGTGTCTTCCAAAGGGGGCCATGTGGACACATCTCCCTGTCCAAATGGCCAC	4080
QY	4081	TTTAAAGAGGAGGGGATGCTCTGAGAAAGCAGCCAGGGGTGAGAGTGCATCATATGTT	4140
Db	4081	TTTAAAGAGGAGGGGATGCTCTGAGAAAGCAGCCAGGGGTGAGAGTGCATCATATGTT	4140
QY	4141	CTGGAAAAACCCCAAGCTTCTCCCTTGGAGATCTCTCTGAAATGATTCATATATGCCCTG	4200
Db	4141	CTGGAAAAACCCCAAGCTTCTCCCTTGGAGATCTCTCTGAAATGATTCATATATGCCCTG	4200
QY	4201	CGCTTCATTCCTCCGTCACCTCTGTGTGTGTGTTTACCAACGCGCTTCATCTTGAGAAATGC	4260
Db	4201	CGCTTCATTCCTCCGTCACCTCTGTGTGTGTGTTTACCAACGCGCTTCATCTTGAGAAATGC	4260
QY	4261	ACCTTCAACCTCTACCTGATCCGAAGATCGTGCATTTGGGAAAGAACTGGAGCTGTGC	4320
Db	4261	ACCTTCAACCTCTACCTGATCCGAAGATCGTGCATTTGGGAAAGAACTGGAGCTGTGC	4320
QY	4321	TATGAAAGCCCTGGAGAAACCAAGCTGTCTCGAGATTCTACGTTGGCCACTTGGGATCA	4380
Db	4321	TATGAAAGCCCTGGAGAAACCAAGCTGTCTCGAGATTCTACGTTGGCCACTTGGGATCA	4380
QY	4381	GGGATCAAGGCTGCAAGTGAAAGCAAGAAAGATAGACCTGCGTGTGGAGAGCCTTGTGTG	4440
Db	4381	GGGATCAAGGCTGCAAGTGAAAGCAAGAAAGATAGACCTGCGTGTGGAGAGCCTTGTGTG	4440
QY	4441	AAACCAAGAGATCTCAATGCTCGGAAATATCTGTGATTCCTCCAGGCCGCAATAGCCGTACTCT	4500
Db	4441	AAACCAAGAGATCTCAATGCTCGGAAATATCTGTGATTCCTCCAGGCCGCAATAGCCGTACTCT	4500
QY	4501	TCACCTCTGATGCCCCGCAAGTTCGTGCACTTTGTGACCAAGTATCGAGAGCAGCTGATA	4560
Db	4501	TCACCTCTGATGCCCCGCAAGTTCGTGCACTTTGTGACCAAGTATCGAGAGCAGCTGATA	4560
QY	4561	GCCCGAGTGACATCGTGGAGAGTGTCTTGGACAAACTGACATGGACAGAGTGTCTGAGCCAG	4620
Db	4561	GCCCGAGTGACATCGTGGAGAGTGTCTTGGACAAACTGACATGGACAGAGTGTCTGAGCCAG	4620
QY	4621	GAGCAGTACGGAAGGGGTGTGGCTGGAACACGAGGCCAGGCCAGATGGCCGAAAGCTTTTC	4680
Db	4621	GAGCAGTACGGAAGGGGTGTGGCTGGAACACGAGGCCAGGCCAGATGGCCGAAAGCTTTTC	4680
QY	4681	AGCTTGAAGCAGTCTGTGGACCGGAAGTGCAAAGATGGAATCTTACCAAGCCTGTGAAGAG	4740
Db	4681	AGCTTGAAGCAGTCTGTGGACCGGAAGTGCAAAGATGGAATCTTACCAAGCCTGTGAAGAG	4740
QY	4741	ACCATCTCTCACTTCATTTATGGAACCTTGGGAAAGGGCAGCAAAAAGGACCTCTTGCCA	4800

Db	4741		ACCATCTCACTCATTATGAACTCTGGAGGAAGGGCAGCAAAAAGGACTCTTGCA	4800
Qy	4801		CTCAGCAGCTGAAGTATCAACACCAGCCCTTGACCTCTGAGTCCTGTGGCTTTGGCTGACCC	4860
Db	4801		CTCAGCAGCTGAAGTATCAACACCAGCCCTTGACCTCTGAGTCCTGTGGCTTTGGCTGACCC	4860
Qy	4861		TTCCTTTGGGTCCTCAGTTTCTTTCTCTCGAACAAGAGTTGCCATCTGTGTTTGCCTTCACGA	4920
Db	4861		TTCCTTTGGGTCCTCAGTTTCTTTCTCTCGAACAAGAGTTGCCATCTGTGTTTGCCTTCACGA	4920
Qy	4921		CTAAGTAAGAAGAACTTGTATGAGCCTTGGCTGGGCACTTAATGTGTCATAGCCAGGAGTAG	4980
Db	4921		CTAAGTAAGAAGAACTTGTATGAGCCTTGGCTGGGCACTTAATGTGTCATAGCCAGGAGTAG	4980
Qy	4981		CCAAGAGGGGCCCCAGTCCAGAGTGGGCTTAAACAGATCTCAGGGAATGCCATCTGAGACT	5040
Db	4981		CCAAGAGGGGCCCCAGTCCAGAGTGGGCTTAAACAGATCTCAGGGAATGCCATCTGAGACT	5040
Qy	5041		GCGAAGAACCCCTGCGACCTCATATGAGCCTTCATCTGTGGTGGCCACAGCACGCCAAGCTTAGA	5100
Db	5041		GCGAAGAACCCCTGCGACCTCATATGAGCCTTCATCTGTGGTGGCCACAGCACGCCAAGCTTAGA	5100
Qy	5101		GCCTCTCCGGATCCCATCTCAGGCGGCAAGAGAAATGAGGGGACATGTGAACCATTTGCCCTC	5160
Db	5101		GCCTCTCCGGATCCCATCTCAGGCGGCAAGAGAAATGAGGGGACATGTGAACCATTTGCCCTC	5160
Qy	5161		TGGCTGTGTGTCACAGGATGAGGCCCCAAAAATTTGGGGTTTCAAGCGTGGAGGGCCAGTGGATTTC	5220
Db	5161		TGGCTGTGTGTGTCACAGGATGAGGCCCCAAAAATTTGGGGTTTCAAGCGTGGAGGGCCAGTGGATTTC	5220
Qy	5221		TTGGCTTTTGTATCAGGAAGATCTACAGAGCAAGCCCAAAGATTAAGTGAAGGAAGTTT	5280
Db	5221		TTGGCTTTTGTATCAGGAAGATCTACAGAGCAAGCCCAAAGATTAAGTGAAGGAAGTTT	5280
Qy	5281		AITCAGAAATAAAGAGATCACAGCTCTTTTGAATTTTGTCTACAGCAGGCTTTCCAGTT	5340
Db	5281		AITCAGAAATAAAGAGATCACAGCTCTTTTGAATTTTGTCTACAGCAGGCTTTCCAGTT	5340
Qy	5341		TTTACCGAAGAAACCCCTATTAATTTTATTTTACTCTTAAATTTAAGAAATTAAGAAAT	5400
Db	5341		TTTACCGAAGAAACCCCTATTAATTTTATTTTACTCTTAAATTTAAGAAATTAAGAAAT	5400
Qy	5401		ACAAAAAAGAAAAATGAAAAATTAAGGAATTAAGAGTTTACTCTAC	5444
Db	5401		ACAAAAAAGAAAAATGAAAAATTAAGGAATTAAGAGTTTACTCTAC	5444
 RESULT 4				
US-09-956-712-11				
; Sequence 11, Application US/09956712				
; Publication No. US20030092648A1				
GENERAL INFORMATION:				
APPLICANT: C. Frank Bennett				
APPLICANT: Susan M. Freiler				
TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION				
FILE REFERENCE: RTS-0326				
CURRENT APPLICATION NUMBER: US/09/956,712				
CURRENT FILING DATE: 2001-09-19				
NUMBER OF SEQ ID NOS: 91				
SEQ ID NO 11				
LENGTH: 6531				
TYPE: DNA				
ORGANISM: Homo sapiens				
FEATURE:				
US-09-956-712-11				
 Query Match 97.1%; Score 5284.2; DB 11; Length 6531;				
Best Local Similarity 97.5%; Pzed. No. 0;				
Matches 5441; Conservative 0; Mismatches 3; Indels 135; Gaps 2;				
Qy	1		GCCCCAGGCGCTGGAGAGGTCTGAGAAACCTTGGAGCCAGCAGCCCGGGGCTCACTCT	60

Dh 1 GCCCGAGGGGCTGAGAGGCTGTGAAGAACTTGAGAGCCAGAGCCCGGGGCTCCACTT 60
Qy 61 GGGTCTGAAGAGCCATTCCTGCTGCTGCGGCTCTCCACCCCACTCTTCTCAGCCTT 120
Db 61 GGGTCTGAAGAGCCATTCCTGCTGCTGCGGCTCTCCACCCCACTCTTCTCAGCCTT 120
Qy 121 GCAGCTCAAGGCTGATCTCAGAGAGTCCAGAGCCAGAGAGGGAGAAATCTGAGAA 180
Db 121 GCAGCTCAAGGCTGATCTCAGAGAGTCCAGAGCCAGAGAGGGAGAAATCTGAGAA 180
Qy 181 CAGAAAGTGAAGCTTGCCGACACCCCATCTCCGTCACCACTCTCCCTCACCCCTCAC 240
Db 181 CAGAAAGTGAAGCTTGCCGACACCCCATCTCCGTCACCACTCTCCCTCACCCCTCAC 240
Qy 241 CCTCCCTGCGGCTGAGAGCCCATCCAGAGCTCCCTATCAGTGAATCTTCTTCCAGT 300
Db 241 CCTCCCTGCGGCTGAGAGCCCATCCAGAGCTCCCTATCAGTGAATCTTCTTCCAGT 300
Qy 301 GTCTTGAAGAGCCCTCTGAGGCTCTCCCTGCTGCTTCTTCTTCACTTCCCTCTAT 360
Db 301 GTCTTGAAGAGCCCTCTGAGGCTCTCCCTGCTGCTTCTTCTTCACTTCCCTCTAT 360
Qy 361 CCGGCTCTATCTGAGGCTGCTGAGATTTATAAAGCTGGGTTCCGAATGCTGAATA 420
Db 361 CCGGCTCTATCTGAGGCTGCTGAGATTTATAAAGCTGGGTTCCGAATGCTGAATA 420
Qy 421 GACGGTAAGAGCCAGAGCAAGACAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 GACGGTAAGAGCCAGAGCAAGACAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 CTGAGAAATCTCCCAAGACACCTCTTAACTCCGAGAGAGAGATGCTGCGAGAGCTG 540
Db 481 CTGAGAAATCTCCCAAGACACCTCTTAACTCCGAGAGAGAGATGCTGCGAGAGCTG 540
Qy 541 GGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 GGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 601 CTGCTGCAATTAAGAGCACTCCAGAGCTCTTCCGAGTGAAGACCCGCTCAGCCAG 660
Db 601 CTGCTGCAATTAAGAGCACTCCAGAGCTCTTCCGAGTGAAGACCCGCTCAGCCAG 660
Qy 661 AAGACGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
Db 661 AAGACGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
Qy 721 TGGGACCTTGAAGCTTCACTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
Db 721 TGGGACCTTGAAGCTTCACTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
Qy 781 CAGGAAAGGAGAGAGCACTCTCCCTCACTCCCTCACTCCCTCACTCCCTCACTCC 840
Db 781 CAGGAAAGGAGAGAGCACTCTCCCTCACTCCCTCACTCCCTCACTCCCTCACTCC 840
Qy 841 TCTTCCAGCCCACTCCCTCACTCCCTCACTCCCTCACTCCCTCACTCCCTCACT 900
Db 841 TCTTCCAGCCCACTCCCTCACTCCCTCACTCCCTCACTCCCTCACTCCCTCACT 900
Qy 901 GGGTGAAG 960
Db 901 GGGTGAAG 960
Qy 961 CGCTGAGAGAAATCTGCTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCACT 1020
Db 961 CGCTGAGAGAAATCTGCTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCACT 1020
Qy 1021 GAGTCTCAAG 1080
Db 1021 GAGTCTCAAG 1080
Qy 1081 GAGTCTCAAG 1140
Db 1081 GAGTCTCAAG 1140

Qy 1141 CCTTGAATGAAG 1200
Db 1141 CCTTGAATGAAG 1200
Qy 1201 TCAG 1260
Db 1201 TCAG 1260
Qy 1261 AGCTTCAAG 1320
Db 1261 AGCTTCAAG 1320
Qy 1321 TGGGCTTGAAG 1380
Db 1321 TGGGCTTGAAG 1380
Qy 1381 CCTTCAAG 1440
Db 1381 CCTTCAAG 1440
Qy 1441 AATCGAG 1500
Db 1441 AATCGAG 1500
Qy 1501 CCTGAG 1560
Db 1501 CCTGAG 1560
Qy 1561 GTGAAG 1620
Db 1561 GTGAAG 1620
Qy 1621 TTGAGCTGAG 1680
Db 1621 TTGAGCTGAG 1680
Qy 1681 GATGAG 1740
Db 1681 GATGAG 1740
Qy 1741 TTGATCTGAG 1800
Db 1741 TTGATCTGAG 1800
Qy 1801 TGTCTGAG 1860
Db 1801 TGTCTGAG 1860
Qy 1861 ACTATACTTCCGAG 1920
Db 1861 ACTATACTTCCGAG 1920
Qy 1921 ATTCTCTTGAAG 1980
Db 1921 ATTCTCTTGAAG 1980
Qy 1981 AAGAGATATTTCAAG 2040
Db 1981 AAGAGATATTTCAAG 2040
Qy 2041 GTCAAAATCAAG 2100
Db 2041 GTCAAAATCAAG 2100
Qy 2101 TGAATCTGAG 2160
Db 2101 TGAATCTGAG 2160
Qy 2161 ACCAG 2220
Db 2161 ACCAG 2220

2221 CAGCTCAGAGACTCTGCTCTCTGAGGAGCATCTGCAAAAAAGACCTTTTC 2280
2221 CAGCTCAGAGACTCTGCTCTCTGAGGAGCATCTGCAAAAAAGACCTTTTC 2280
2281 AGTCAGATGACCTCAGAGAGCATGGATTAGATGGGGCATCATCTCCACTTTCTTGAAG 2340
2281 AGTCAGATGACCTCAGAGAGCATGGATTAGATGGGGCATCATCTCCACTTTCTTGAAG 2340
2341 ATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTACAGCTTCACTCTGTTTC 2400
2341 ATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTACAGCTTCACTCTGTTTC 2400
2401 CAAGAGTTCTTTGAGAGCAATGCTCTGAGAGATGAGAGAGAGAGTAAACAT 2460
2401 CAAGAGTTCTTTGAGAGATGCTCTGATCTTGGAGATGAGAGAGAGAGTAAACAT 2460
2461 TCTAATTGATCATATAGATTGAGAAAGACGCTAGAGATATGAGATATCATGGCTGTTT 2520
2461 TCTAATTGATCATATAGATTGAGAAAGACGCTAGAGATATGAGATATCATGGCTGTTT 2520
2521 GGGGATCAACCAACGTTTCTTATTTGGGCTGTTAAGTATGAGAGAGAGAGATG 2580
2521 GGGGATCAACCAACGTTTCTTATTTGGGCTGTTAAGTATGAGAGAGAGAGATG 2580
2581 GAGAAACATCTTCACTGCGGCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGATG 2640
2581 GAGAAACATCTTCACTGCGGCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGATG 2640
2641 CTGAGAGCTGCTGCTGACAGCACTCTCTGAGAGTCCCTCACTGCTTGTACAGAGATCG 2700
2641 CTGAGAGCTGCTGCTGACAGCACTCTCTGAGAGTCCCTCACTGCTTGTACAGAGATCG 2700
2701 AACAAACGTTCTCTGACACAGATGAGGCCATTTGAGAAATGGGATGTTGTAGAA 2760
2701 AACAAACGTTCTCTGACACAGATGAGGCCATTTGAGAAATGGGATGTTGTAGAA 2760
2761 AACAGATGAGAGCTCTTATGATGACCTTTCGATTAATTCAGCGGCAAGTGAAGAG 2820
2761 AACAGATGAGAGCTCTTATGATGACCTTTCGATTAATTCAGCGGCAAGTGAAGAG 2820
2821 CTTACAGTATTTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATGCTG 2880
2821 CTTACAGTATTTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATGCTG 2880
2881 TTCAAGTGGGCTCCAGTCAAGATGCTATTTGGAGATTTCTTCTCCGCTCTCAAGGTC 2940
2881 TTCAAGTGGGCTCCAGTCAAGATGCTATTTGGAGATTTCTTCTCCGCTCTCAAGGTC 2940
2941 ACCAGAAACCTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3000
2941 ACCAGAAACCTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3000
3001 AGCTTTTAAAGACCTCTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3060
3001 AGCTTTTAAAGACCTCTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3060
3061 TGTGGCTCTCAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3120
3061 TGTGGCTCTCAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3120
3121 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3180
3121 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3180
3181 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3240
3181 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3240
3241 ACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3300
3241 ACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3300
3301 CTAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3360
3301 CTAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGAGATG 3360

3301 CTAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3360
3361 AGGATCTCTGCTGCAAACTCATATGCTCTGAGAGAGAGAGAGAGAGATGAG 3420
3361 AGGATCTCTGCTGCAAACTCATATGCTCTGAGAGAGAGAGAGAGAGATGAG 3420
3421 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3480
3421 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGAGATG 3480
3481 CGAGAAACCAAGTATGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3540
3481 CGAGAAACCAAGTATGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3540
3541 ACATCTCTCACTCAAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3600
3541 ACATCTCTCACTCAAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3600
3601 GCTAATCTCAAACTCTGAGAGAGAGAGAGAGATCAATCATGAGAGAGATG 3660
3601 GCTAATCTCAAACTCTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3660
3658 GAAAGCTCTCCAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3717
3658 GAAAGCTCTCCAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3717
3661 GAAAGCTCTCCAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3720
3661 GAAAGCTCTCCAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3720
3718 GGGAGAGCTGATCAAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3777
3718 GGGAGAGCTGATCAAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3777
3721 GGGAGAGCTGATCAAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3780
3721 GGGAGAGCTGATCAAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3780
3778 GTGGCTACTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3837
3778 GTGGCTACTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3837
3781 GTGGCTACTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3840
3781 GTGGCTACTGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3840
3838 GGTCTCTCAACGCTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3897
3838 GGTCTCTCAACGCTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3897
3841 GGTCTCTCAACGCTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3900
3841 GGTCTCTCAACGCTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3900
3898 GAGATGAGATCTGATGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3957
3898 GAGATGAGATCTGATGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3957
3901 GAGATGAGATCTGATGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3960
3901 GAGATGAGATCTGATGAGAGAGAGAGATCAATCATGAGAGAGAGATG 3960
3958 ATGATGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4017
3958 ATGATGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4017
3961 ATGATGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4020
3961 ATGATGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4020
4018 CTCCCTCACTTTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4077
4018 CTCCCTCACTTTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4077
4021 CTCCCTCACTTTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4080
4021 CTCCCTCACTTTGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4080
4078 CACTTTAAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4137
4078 CACTTTAAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4137
4081 CACTTTAAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4140
4081 CACTTTAAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4140
4138 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4197
4138 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4197
4141 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4200
4141 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4200
4198 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4257
4198 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4257
4201 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4260
4201 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4260
4258 GTCACTTCAACCTTCACTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4302
4258 GTCACTTCAACCTTCACTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4302
4261 GTCACTTCAACCTTCACTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4320
4261 GTCACTTCAACCTTCACTGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4320
4303 ----- 4302
4303 ----- 4302
4321 CTAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4380
4321 CTAGAGAGAGAGAGAGAGAGAGAGAGAGATCAATCATGAGAGAGAGATG 4380
4303 -----AAG 4305
4303 -----AAG 4305

Db 4381 ATGGGCTGTCTTACACTGTGTCTGGGTCTGTTCAAGGATGCTGAATACTCCCAAG 4440
 Qy 4306 GAACCTGAGCTCTGTATCGAAGCCCTGGAAGAACAGTGTCTCGAGTTCTACGTT 4365
 Db 4441 GAACTGAGCTTGTCTATCGAAGCCCTGGAAGAACAGTGTCTCGAGTTCTACGTT 4500
 Qy 4366 GGCCTATTGGGATCAGGATCAGGCTCGCAAGTGAAGACAGAAATGAGACTCTGGTG 4425
 Db 4501 GGCCTATTGGGATCAGGATCAGGCTCGCAAGTGAAGACAGAAATGAGACTCTGGTG 4560
 Qy 4426 TGGGAGGCTTGGTGAACACAGAGATCTCATGCTGCAACTCTGATCCCTCCAGCC 4485
 Db 4561 TGGGAGGCTTGGTGAACACAGAGATCTCATGCTGCAACTCTGATCCCTCCAGCC 4620
 Qy 4486 CGCATAGCCGTACTTCACTCTGATGTCGCCAGTTGCTGCACTTTGACACGAT 4545
 Db 4621 CGCATAGCCGTACTTCACTCTGATGTCGCCAGTTGCTGCACTTTGACACGAT 4680
 Qy 4546 CGAGAGCAGTGTATACCCGAGTGAATCGGTGGAAGTTGTCTTGGACAACTGCAATGA 4605
 Db 4681 CGAGAGCAGTGTATACCCGAGTGAATCGGTGGAAGTTGTCTTGGACAACTGCAATGA 4740
 Qy 4606 CAGGTCTGAGCAGGAGCAGTACGAGAGGGTGTGGCTGAGAAACAGAGCCACGAC 4665
 Db 4741 CAGGTCTGAGCAGGAGCAGTACGAGAGGGTGTGGCTGAGAAACAGAGCCACGAC 4800
 Qy 4666 ATGCGGAGCTGTCTTCACTTGAAGCAGTCTCTGGAACCGAAGTGAAGTGAATCTTAC 4725
 Db 4801 ATGCGGAGCTGTCTTCACTTGAAGCAGTCTCTGGAACCGAAGTGAAGTGAATCTTAC 4860
 Qy 4726 CAGCCCTGAAGAGACCCATCTCTCACTTATATGAATCTTGGGAGAAAGGCGACAA 4785
 Db 4861 CAGCCCTGAAGAGACCCATCTCTCACTTATATGAATCTTGGGAGAAAGGCGACAA 4920
 Qy 4786 AAGGAGCTCTGCACTGAGCAGTGAAGTCAACACGACCCCTGACCTTGAAGTCT 4845
 Db 4921 AAGGAGCTCTGCACTGAGCAGTGAAGTCAACACGACCCCTGACCTTGAAGTCT 4980
 Qy 4846 GGGTTGGCTGACCCCTTCTTGGGTCTCAATTTCTTCTGCAAAACAAGTTCCATCTG 4905
 Db 4981 GGGTTGGCTGACCCCTTCTTGGGTCTCAATTTCTTCTGCAAAACAAGTTCCATCTG 5040
 Qy 4906 GTTGGCTTCCAGCACTTAAAGTATGGAATTTGATGATGCTTGTCTGGGCACTTATG 4965
 Db 5041 GTTGGCTTCCAGCACTTAAAGTATGGAATTTGATGATGCTTGTCTGGGCACTTATG 5100
 Qy 4966 TCCATCCAGGAGTGCACAGGGGGCCCAAGTCAAGTGGCTTAACAGCATCTCAGGAA 5025
 Db 5101 TCCATCCAGGAGTGCACAGGGGGCCCAAGTCAAGTGGCTTAACAGCATCTCAGGAA 5160
 Qy 5026 TGTTCATCTGAGCTGGCAAGACCCCTGCAAGCTCATAGAGCTCATCTGTGTGCACA 5085
 Db 5161 TGTTCATCTGAGCTGGCAAGACCCCTGCAAGCTCATAGAGCTCATCTGTGTGCACA 5220
 Qy 5086 GCGAGCCAGCTTGAAGACCCCTCCGGAATCCCATCAAGGCGCAAAAGAAATGAGGGAAT 5145
 Db 5221 GCGAGCCAGCTTGAAGACCCCTCCGGAATCCCATCAAGGCGCAAAAGAAATGAGGGAAT 5280
 Qy 5146 GGAACCATTTGCTCTGAGCTGTGTCAAGGGGTGAGCCCAAAATTTGGGGTTGAGCTGGG 5205
 Db 5281 GGAACCATTTGCTCTGAGCTGTGTCAAGGGGTGAGCCCAAAATTTGGGGTTGAGCTGGG 5340
 Qy 5206 AGGCCAGCTGTGATTTTGGCTTTGTACAGGAAGATCTAACAGACCAAGCAAGATTA 5265
 Db 5341 AGGCCAGCTGTGATTTTGGCTTTGTACAGGAAGATCTAACAGACCAAGCAAGATTA 5400
 Qy 5266 AGTGAAGAGAGTTTATTCAGAAATTAAGAGATCAAGAGCTTTTAAATTTGCTA 5325
 Db 5401 AGTGAAGAGAGTTTATTCAGAAATTAAGAGATCAAGAGCTTTTAAATTTGCTA 5460
 Qy 5326 GCGAGGCTTTCAGATTTTACAGAAACCCCTATTAATTTAAATTTTACTTAATTT 5385
 Db 5461 GCGAGGCTTTCAGATTTTACAGAAACCCCTATTAATTTAAATTTTACTTAATTT 5520

Qy 5386 AAGAAATTAATAAATACAAAAAGAAAAATGAAAAATTAAGAAATTAAGATTACTTAC 5444
 Db 5521 AAGAAATTAATAAATTAATAAATAAGAAAAATGAAAAATTAAGAAATTAAGATTACTTAC 5579
 RESULT 5
 US-09-956-712-12
 ; Sequence 12, Application US/09956712
 ; Publication No. US20030092648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan M. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
 ; FILE REFERENCE: RTS-0326
 ; CURRENT APPLICATION NUMBER: US/09/956,712
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 12
 ; LENGTH: 5100
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (160)...(4449)
 US-09-956-712-12
 Query Match 89.7%; Score 4882.4; DB 11; Length 5100;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 Qy 516 GACAGAGATGCTGGCGGAGCTTGGCGGAGCTTGGCTGCTTGAAGTTCTGGAAGA 575
 Db 153 GAACAAGATGCTGGCGGAGCTTGGCGGAGCTTGGCTGCTTGAAGTTCTGGAAGA 212
 Qy 576 GAGAGAGTGAAGAGATTCAGCTTCTGCTGCAATTAAGGCACTCCAGAGCTTCT 635
 Db 213 GAGAGAGTGAAGAGATTCAGCTTCTGCTGCAATTAAGGCACTCCAGAGCTTCT 272
 Qy 636 GGGTGAACACCCGCTCAGCCAGAGAGACGATGAGTGAAGTGGCTCGTACCTGTG 695
 Db 273 GGGTGAACACCCGCTCAGCCAGAGAGACGATGAGTGAAGTGGCTCGTACCTGTG 332
 Qy 696 GGGTCAATATGGGAGACAGCGGGCTTGGGACCTAGCCCTTCAATCTTGGGAGCAATGG 755
 Db 333 GGGTCAATATGGGAGACAGCGGGCTTGGGACCTAGCCCTTCAATCTTGGGAGCAATGG 392
 Qy 756 GGTGAGTCACTGTGGCCCAAGCCAGGAGGGGCAAGGCACTCTCCCTCAATCCCTA 815
 Db 393 GGTGAGTCACTGTGGCCCAAGCCAGGAGGGGCAAGGCACTCTCCCTCAATCCCTA 452
 Qy 816 CAGCCCAAGTGAACCCCACTGGGGTCTTCCAGCCCAACCACTTCCAGCCGAGTGTAA 875
 Db 453 CAGCCCAAGTGAACCCCACTGGGGTCTTCCAGCCCAACCACTTCCAGCCGAGTGTAA 512
 Qy 876 GGCCTGAGATCAATGAATTCGGCGGGGTGACACCCAGGGCTCAAGAGAAAGGTTTGA 935
 Db 513 GGCCTGAGATCAATGAATTCGGCGGGGTGACACCCAGGGCTCAAGAGAAAGGTTTGA 572
 Qy 936 ACAGCTGCTGACATCTGAGCGCGCTGAGAGAAATCTTGTGCTCACTCCTTACCA 995
 Db 573 ACAGCTGCTGACATCTGAGCGCGCTGAGAGAAATCTTGTGCTCACTCCTTACCA 632
 Qy 996 AGCTCTTCCAGCTCCCAAGACCATAGTCTCCAGCCAGAGATCAACCAAGCCCCAC 1055
 Db 633 AGCTCTTCCAGCTCCCAAGACCATAGTCTCCAGCCAGAGATCAACCAAGCCCCAC 692
 Qy 1056 ATCCACAGAGTGTGGGAGCTGGGAGTCCCACTTCAAGCCCAAGCTTCAAGCA 1115
 Db 693 ATCCACAGAGTGTGGGAGCTGGGAGTCCCACTTCAAGCCCAAGCTTCAAGCA 752
 Qy 1116 GCAGAGGCTCTCGGAGCCCAATGAGCTTGTGATGAAAGTGAATTTACTTACAGAG 1175

Db 753 GCAGAGGCTCCTGGACCAATGGCTCTGGATGAACGTCAAGAAATTACTACAGA 812
Qy 1176 AATCAGAGAAAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGCAGGGGTGT 1235
Db 813 AATCAGAGAAAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGCAGGGGTGT 872
Qy 1236 AGGAAGCCCCCAGAGGCGCACACAGCTACAGGCCCAACAGCCCATGGAGCCCTC 1295
Db 873 AGGAAGCCCCCAGAGGCGCACACAGCTACAGGCCCAACAGCCCATGGAGCCCTC 932
Qy 1296 TGTGAGAGAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGATTTTAAACAAAATT 1355
Db 933 TGTGAGAGAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGATTTTAAACAAAATT 992
Qy 1356 CACACAGCTGTACTTCTTCAAAAGCTCACCCCAAGAACCAAGATCCCTGTCTAGAG 1415
Db 993 CACACAGCTGTACTTCTTCAAAAGCTCACCCCAAGAACCAAGATCCCTGTCTAGAG 1052
Qy 1416 AAGCTGGCTGTATATGTGAGAGAAATCGAGGACATTTAATTGAGATCAGAGCTTATT 1475
Db 1053 AAGCTGGCTGTATATGTGAGAGAAATCGAGGACATTTAATTGAGATCAGAGCTTATT 1112
Qy 1476 TGGCCCAAGGCTTGATACCAAGAACTCGATAGTCACTAGAGGGGCTGTGTGAT 1535
Db 1113 TGGCCCAAGGCTTGATACCAAGAACTCGATAGTCACTAGAGGGGCTGTGTGAT 1172
Qy 1536 TGGGAAGTCAACACTGGCCAGGCGAGGTGAAGAAAGCTGGGGGAGAGGCGACTGTATG 1595
Db 1173 TGGGAAGTCAACACTGGCCAGGCGAGGTGAAGAAAGCTGGGGGAGAGGCGACTGTATG 1232
Qy 1596 GGAACGGCTTCAGAGATGTCTTCTTCACTTCAAGTGCAGAGAGCTGGCCCATGAGTGTGT 1655
Db 1233 GGAACGGCTTCAGAGATGTCTTCTTCACTTCAAGTGCAGAGAGCTGGCCCATGAGTGTGT 1292
Qy 1656 GAGTCTGCTGAGCTACCGAGAAAGATGGAGCAGCCACTCCGGCTCCCATTAAGACAGAT 1715
Db 1293 GAGTCTGCTGAGCTACCGAGAAAGATGGAGCAGCCACTCCGGCTCCCATTAAGACAGAT 1352
Qy 1716 CTTGTCTTAGGCGAAGCGGGCTCTTCTTCACTCTGATGTGTAGATGAAGCAGGATGGT 1775
Db 1353 CTTGTCTTAGGCGAAGCGGGCTCTTCTTCACTCTGATGTGTAGATGAAGCAGGATGGT 1412
Qy 1776 CTTGCAAGAGCGGAGTCTGAGCTCTGTCTGCACTGAGCGAGCCAGCGGGGAGTGC 1835
Db 1413 CTTGCAAGAGCGGAGTCTGAGCTCTGTCTGCACTGAGCGAGCCAGCGGGGAGTGC 1472
Qy 1836 ACTGCTGGGCGAGTTGCTGGGAGAACTATACTTCCCGAGGCTCTTCTGATCAAGGC 1895
Db 1473 ACTGCTGGGCGAGTTGCTGGGAGAACTATACTTCCCGAGGCTCTTCTGATCAAGGC 1532
Qy 1896 TCGGACCAAGCTCTGAGAACTCTTCTTCTTGGAGGAGCAGTTGGGTAGAGT 1955
Db 1533 TCGGACCAAGCTCTGAGAACTCTTCTTCTTGGAGGAGCAGTTGGGTAGAGT 1592
Qy 1956 CCTGGGGTCTCTGAGTCCAGAGAGAAAGAAATATTTCTACAGAAATTTCAAGATGAAAG 2015
Db 1593 CCTGGGGTCTCTGAGTCCAGAGAGAAAGAAATATTTCTACAGAAATTTCAAGATGAAAG 1652
Qy 2016 GCAAGCAATTAGAGCTTTAGTGTGTCAATCAAAAGAGCTCTGGGCGCTGTGTCT 2075
Db 1653 GCAAGCAATTAGAGCTTTAGTGTGTGTCAATCAAAAGAGCTCTGGGCGCTGTGTCT 1712
Qy 2076 TGTGCGCTGGGTGTCTGTGCTGGCTGTGACTTGGCTGTGATGAGCAGATGAAGCGGAAGA 2135
Db 1713 TGTGCGCTGGGTGTCTGTGCTGGCTGTGACTTGGCTGTGATGAGCAGATGAAGCGGAAGA 1772
Qy 2136 AAAAATCAGACTGACTTTCAGAGACCAAGCAACCTCTGTCTAATTAATTTGCGCCAGGC 2195
Db 1773 AAAAATCAGACTGACTTTCAGAGACCAAGCAACCTCTGTCTAATTAATTTGCGCCAGGC 1832
Qy 2196 TCTCAAGCTCAGGCAATTGGAGCCCAAGCTCAGAGACTCTGTCTCTGTGGCTGTGAGGG 2255
Db 1833 TCTCAAGCTCAGGCAATTGGAGCCCAAGCTCAGAGACTCTGTCTCTGTGGCTGTGAGGG 1892

Qy 2256 CATCTGGCAAAAAAGACCCCTTTTCACTCCAGATGACCTCAGAGAGCATGGGTAGATGG 2315
Db 1893 CATCTGGCAAAAAAGACCCCTTTTCACTCCAGATGACCTCAGAGAGCATGGGTAGATGG 1952
Qy 2316 GGCCATCATCTCCACTTCTTGAAGATGGGTATCTTCAAGAGCAGCCCATCCCTGTAG 2375
Db 1953 GGCCATCATCTCCACTTCTTGAAGATGGGTATCTTCAAGAGCAGCCCATCCCTGTAG 2012
Qy 2376 CTACAGCTTCACTTCACTCTGTCTTCAAGAGTCTTTCAGAGAAATGTCTATGTCTTGA 2435
Db 2013 CTACAGCTTCACTTCACTCTGTCTTTCAGAGATCTTTCAGAGCAATGTCTATGTCTTGA 2072
Qy 2436 GGAATGAAGGGGAGAGCTTAACTTTAATGTGATATAGATTTGGAAAAAGAGCTTGA 2495
Db 2073 GGAATGAAGGGGAGAGCTTAACTTTAATGTGATATAGATTTGGAAAAAGAGCTTGA 2132
Qy 2496 AGCATATGAATATACATGGCCGTGTGTGGGGCATCAACACAGTTTCTATTTGGGCTGT 2555
Db 2133 AGCATATGAATATACATGGCCGTGTGTGGGGCATCAACACAGTTTCTATTTGGGCTGT 2192
Qy 2556 AAGTATGAAGGGGAGAGAGATGAGAAATCTTTTCACTGCCGCTGTCTCAGGGAG 2615
Db 2193 AAGTATGAAGGGGAGAGAGATGAGAAATCTTTTCACTGCCGCTGTCTCAGGGAG 2252
Qy 2616 GAACCTGATGAGTGGTCCCGTCCCTGCACTGCTGTGTGAGCACACTCTGTGAGTC 2675
Db 2253 GAACCTGATGAGTGGTCCCGTCCCTGCACTGCTGTGTGAGCACACTCTGTGAGTC 2312
Qy 2676 CCTTCACGTGTGATGAGAGCTGGGAAACAAAGCTTCTGTGACAAGATGAGGCCATT 2735
Db 2313 CCTTCACGTGTGTATGAGAGCTGGGAAACAAAGCTTCTGTGACAAGATGAGGCCATT 2372
Qy 2736 CGAAGAAATGGGCAATGTGTGAGAAACAGACATGAGCTCTTATGTGCACTTTCTGAT 2795
Db 2373 CGAAGAAATGGGCAATGTGTGAGAAACAGACATGAGCTCTTATGTGCACTTTCTGAT 2432
Qy 2796 TAAATTAAGCCCGCAGCTGAAGAGCTTCACTGATTTGAGGGCAGGCAACATCAAC 2855
Db 2433 TAAATTAAGCCCGCAGCTGAAGAGCTTCACTGATTTGAGGGCAGGCAACATCAAC 2492
Qy 2856 ATGGAGCCCAACATGATGAGTCCGTTCAGGTGGGTCAGAGTGCAGATGCTTATGGCA 2915
Db 2493 ATGGAGCCCAACATGATGAGTCCGTTCAGGTGGGTCAGAGTGCAGATGCTTATGGCA 2552
Qy 2916 GATTCTCTCTCGCTCCTCAAGTCAACAGAACTGAAGAGCTGAGCTTAATGAGAA 2975
Db 2553 GATTCTCTCTCGCTCCTCAAGTCAACAGAACTGAAGAGCTGAGCTTAATGAGAA 2612
Qy 2976 CTGCTGAGCCACTCTGCAAGTGAAGTCTTTGTAGAACCCTGAGAGCCCTCGCTGCT 3035
Db 2613 CTGCTGAGCCACTCTGCAAGTGAAGTCTTTGTAGAACCCTGAGAGCCCTCGCTGCT 2672
Qy 3036 CCTGGAGACCCCTGGGTGTGGCTGTGGCTGAGCTGAGGACTGCAAGAGACTTTC 3095
Db 2673 CCTGGAGACCCCTGGGTGTGGCTGTGGCTGAGCTGAGGACTGCAAGAGACTTTC 2732
Qy 3096 CTTTGGCTGAGAGCCAAACAGACCTTGAACCTGAGCTGAGCTTGAATGTGTCTAC 3155
Db 2733 CTTTGGCTGAGAGCCAAACAGACCTTGAACCTGAGCTGAGCTTGAATGTGTCTAC 2792
Qy 3156 GGAATGCTGAGAGCCAAACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGCG 3215
Db 2793 GGAATGCTGAGAGCCAAACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGCG 2852
Qy 3216 ACTGCACTGTGATCAGCTGTGGCTCAGGTGTGACTGTGCAAGAGACCTGGCTCTGTGTCT 3275
Db 2853 ACTGCACTGTGATCAGCTGTGGCTCAGGTGTGACTGTGCAAGAGACCTGGCTCTGTGTCT 2912
Qy 3276 TAGTCCAGGCCCAAGCTTGAAGAGCTTGAACCTGCAAGAAACAACTGTGATGAGCTTGG 3335
Db 2913 TAGTCCAGGCCCAAGCTTGAAGAGCTTGAACCTGCAAGAAACAACTGTGATGAGCTTGG 2972

QY	3336	CGTGCACGTGCTGTGAGGGGGCTCAGGCAATCTCGCTGGCAAACTCATATGACCTGGGGCT	3395
Db	2973	CGTGCACGTGCTGTGTAGGGGGCTCAGGCAATCTGCTGGCAAACTCATATGACCTGGGGCT	3032
QY	3396	GGACCAAGACAATCTGTAGTGTATGATGATGAGGACAGAACTGAGGGCCCTGGAGCAGAGAA	3455
Db	3033	GGACCAAGACAATCTGTAGTGTATGATGATGAGGACAGAACTGAGGGCCCTGGAGCAGAGAA	3092
QY	3456	ACCTGACGTGCTATCTTACGACAAACGAAACAAAGTGTATGACCTCTAGGGGCT	3515
Db	3093	GCTTACGTGCTATCTTACGACAAACGAAACAAAGTGTATGACCTCTATTTAGGGGCT	3152
QY	3516	GGATACGGAGAGATGATATATACATCACTCACTCAAGCGGACAGACCTCGATCAGA	3575
Db	3153	GGATACGGAGAGATGATATATACATCACTCACTCAAGCGGACAGACCTCGATCAGA	3212
QY	3576	GAGGGCGGCTTCCATGTTTGCTCAGGCTATCTCAAATCTCTGGAAGTGAAGATCTT	3635
Db	3213	GAGGGCGGCTTCCATGTTTGCTCAGGCTATCTCAAATCTCTGGAAGTGAAGATCTT	3272
QY	3636	CCCAATTGCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACCTTGTG	3695
Db	3273	CCCAATTGCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACCTTGTG	3332
QY	3696	CGTGCCTTCTCTGCTCTCAAGGAGACCTTGACATACGAAGCTTTGGGACCTGACATGA	3755
Db	3333	CGTGCCTTCTCTGCTCTCAAGGAGACCTTGACATACGAAGCTTTGGGACCTGACATGA	3392
QY	3756	CTTCTGGGGCCCCACGGGGCCCTGAGCTACAGATGATGTGACAAAGAAAGAACTTGTG	3815
Db	3393	CTTCTGGGGCCCCACGGGGCCCTGAGCTACAGATGATGTGACAAAGAAAGAACTTGTG	3452
QY	3816	CCGAGTTCACTTCCCTGTAGCTGACTCTACCGCTGAGCCCAACAGGGCTCTCTGCTTGT	3875
Db	3453	CCGAGTTCACTTCCCTGTAGCTGACTCTACCGCTGAGCCCAACAGGGCTCTCTGCTTGT	3512
QY	3876	GATACAGAGAAACGGTGAACCTTGATGATTGAATTTCTGTGTGTGGACCAAGTTCTTGGGTGA	3935
Db	3513	GGTAGAGAAACGGTGAACCGTTGAGATTGAATTTCTGTGTGTGGACCAAGTTCTTGGGTGA	3572
QY	3936	GATCAACCAACAGACAGACTGATGTGTGGACGGGCTCTGCTGACATCAAGCTGAGCC	3995
Db	3573	GATCAACCAACAGACAGACTGATGTGTGGACGGGCTCTGCTGACATCAAGCTGAGCC	3632
QY	3996	TGAGCTGTGSAAGCTGTGCACTTCCCTCACTTTGTGTGCTCTCAAGGGGGCAATGTGA	4055
Db	3633	TGAGCTGTGSAAGCTGTGCACTTCCCTCACTTTGTGTGCTCTCAAGGGGGCAATGTGA	3692
QY	4056	CACATCCCTGTTCCAAATAGGCGCACTTAAAGAGAGGGATGTGCTCGAGAAAGCCAGC	4115
Db	3693	CACATCCCTGTTCCAAATAGGCGCACTTAAAGAGAGGGATGTGCTCGAGAAAGCCAGC	3752
QY	4116	CAGGGTGAAGCTGATCACTATGTTCTGGAAAAACCCAGCTTCTCCCTTGGAGTCT	4175
Db	3753	CAGGGTGAAGCTGATCACTATGTTCTGGAAAAACCCAGCTTCTCCCTTGGAGTCT	3812
QY	4176	CCTGAAAAATGATCCATATATGCCCCGCTTATTTCCGTCACTCTGTGTGTGCTTTA	4235
Db	3813	CCTGAAAAATGATCCATATATGCCCCGCTTATTTCCGTCACTCTGTGTGTGCTTTA	3872
QY	4236	CCACCGGCTCATCTGTAGAAAGTCACTTCACTCACTCACTGATCCCAATGACTGCTC	4295
Db	3873	CCACCGGCTCATCTGTAGAAAGTCACTTCACTCACTCACTGATCCCAATGACTGCTC	3932
QY	4296	CATTTCGGAAGAACTGAGCTCTGCTATGAAAGCCCTGGAAGAACCAAGCTGTTCTCGGA	4355
Db	3933	CATTTCGGAAGAACTGAGCTCTGCTATGAAAGCCCTGGAAGAACCAAGCTGTTCTCGGA	3992
QY	4356	GTTCTACGTTGGCACTTGGGATCAGGGATCAGGCTGTGAATGTAAAGCAAGAAAGATGA	4415
Db	3993	GTTCTACGTTGGCACTTGGGATCAGGGATCAGGCTGTGAATGTAAAGCAAGAAAGATGA	4052
QY	4416	GACTTCGTGTGGAGGCTTGTGTAAACCAAGAGATCTCATGCTCTCACTCTGAT	4475

Db	4053	GACCTCGTGTGGGAGGCCCTTGGTGAACCAAGAGATCTCATGCTGCACACTCTGAT	4112
QY	4476	CCCTCCAGCCCGCATAGCCGTAACCTTCACTCTGGAATGCCCGCAGTTGCTGCACCTTTGT	4535
Db	4113	CCCTCCAGCCCGCATACCCGTAACCTTCACTCTGGAATGCCCGCAGTTGCTGCACCTTTGT	4172
QY	4536	GGACCAAGTATCGAAGAGAGCTGATAGGCCCGAGTACATCCGTTGAGAGTTGCTTGGACAA	4595
Db	4173	GGACCAAGTATCGAAGAGAGCTGATAGGCCCGAGTACATCCGTTGAGAGTTGCTTGGACAA	4232
QY	4596	ACTGCATGAGCAGGTCGTGAGCCAGAGAGCATGAGAGGGTGTGCTGAGAACACGAG	4655
Db	4233	ACTGCATGAGCAGGTCGTGAGCCAGAGAGCATGAGAGGGTGTGCTGAGAACACGAG	4232
QY	4656	GCCGAGCCAGATGGGGAGAGCTGTCACTTGAAGCAGTCCCTGGGAGCCGGAAGTGCAGAA	4715
Db	4293	GCCGAGCCAGATGGGGAGAGCTGTCACTTGAAGCAGTCCCTGGGAGCCGGAAGTGCAGAA	4352
QY	4716	TGCACTCTACCAAGCCCTGGAAGAGACCATCTCCACTTATATGGAACCTTGGAGAA	4775
Db	4353	TGCACTCTACCAAGCCCTGGAAGAGACCATCTCCACTTATATGGAACCTTGGAGAA	4412
QY	4776	GGGCGACGAAAAAGGAGATCTCTGCGCATCTGACAGCTGAAGTATCAACACAGCCCTTGAAC	4835
Db	4413	GGGCGACGAAAAAGGAGATCTCTGCGCATCTGACAGCTGAAGTATCAACACAGCCCTTGAAC	4472
QY	4836	CTTGAGTCTTGCTGGCTTGGAGTGAACCCCTCTTGGAGTCAAGTTTCTTCTCTGCAGAACAG	4895
Db	4473	CTTGAGTCTTGCTGGCTTGGAGTGAACCCCTCTTGGAGTCAAGTTTCTTCTCTGCAGAACAG	4532
QY	4896	TTGCCATCTGGTTTGGCTTTCAGACACTTAAGTATGGAACCTTGAATGATGCTTTGCTGG	4955
Db	4533	TTGCCATCTGGTTTGGCTTTCAGACACTTAAGTATGGAACCTTGAATGATGCTTTGCTGG	4592
QY	4956	GCATTATGTGTCCATGTCGACAGGAGTGCACAGGGGGCCCAAGTCCAGGTGAGCTTAACAGA	5015
Db	4593	GCATTATGTGTCCATGTCGACAGGAGTGCACAGGGGGCCCAAGTCCAGGTGAGCTTAACAGA	4652
QY	5016	TCTCAGGGAAATGTCATCTGTGAGCTGTGCAGAACCCCTGACAGACTCATAGAGCTCATCT	5075
Db	4653	TCTCAGGGAAATGTCATCTGTGAGCTGTGCAGAACCCCTGACAGACTCATAGAGCTCATCT	4712
QY	5076	GATGGCCACAGCAGCAAGCCTTAGAGCCCTCGGATCCCATCCAGGGCGCAAGAGAGATA	5135
Db	4713	GATGGCCACAGCAGCAAGCCTTAGAGCCCTCGGATCCCATCCAGGGCGCAAGAGAGATA	4772
QY	5136	GGAAGGACATGGAACCAATTTGCTCTGTGGCTGTGCACAGGGTGAAGCCCAAAATTTGGGCT	5195
Db	4773	GGAAGGACATGGAACCAATTTGCTCTGTGGCTGTGCACAGGGTGAAGCCCAAAATTTGGGCT	4832
QY	5196	TCACGCGTGGAGGGCCACGCTGATCTCTGGCTTGTGTAAGAAAGATCTACAGAGACCAAGCC	5255
Db	4833	TCACGCGTGGAGGGCCACGCTGATCTCTGGCTTGTGTAAGAAAGATCTACAGAGACCAAGCC	4892
QY	5256	AACAGAGTAAAGTGGAGAGAGTTTATTCAGAAATATAAGAGATATCAAGCTCTTTAG	5315
Db	4893	AACAGAGTAAAGTGGAGAGAGTTTATTCAGAAATATAAGAGATATCAAGCTCTTTAG	4952
QY	5316	AATTTGTCTAGCAGGCTTTCAGATTTTATTCAGAAAAACCTTATATAATTTAAATTTT	5375
Db	4953	AATTTGTCTAGCAGGCTTTCAGATTTTATTCAGAAAAACCTTATATAATTTAAATTTT	5012
QY	5376	ACTTAAATTTAGATTTAAAAAATATACAAAAAAGAAAAATGAAATATAAGAAATAAGAA	5435
Db	5013	ACTTAAATTTAGATTTAAAAAATATACAAAAAAGAAAAATGAAATATAAGAAATAAGAA	5072
QY	5436	GTTA 5439	
Db	5073	GTTA 5076	

RESULT 6

Db 1741 CAAAAAAGACCTTTTCAGTCCAGATGACCTGAGAGCATGGGTTAAGTGGGGCCATC 1800
Qy 2323 ATTCACACTTCTTGAAGATGGGTATCTTCAAGAGCACTCCCTCTGACCTACAGC 2382
Db 1801 ATCTCCACCTTCTTGAAGATGGGTATCTTCAAGAGCACTCCCTCTGACCTACAGC 1860
Qy 2383 TTCAATCACTGTTTCCAGAGTCTTTCAGCAATGTCATATGCTTGGAGATGAG 2442
Db 1861 TTCAATCACTGTTTCCAGAGTCTTTCAGCAATGTCATATGCTTGGAGATGAG 1920
Qy 2443 AAGGGAGAGGTAAACATTTAATTCATATGATATTTGAAAAAGACGTAGAGCATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTAATTCATATGATATTTGAAAAAGACGTAGAGCATAT 1980
Qy 2503 GGAATACATGGCTGTTTGGGGCATCAACACAGTTTCTATTGGGCTGTTAAGTAT 2562
Db 1981 GGAATACATGGCTGTTTGGGGCATCAACACAGTTTCTATTGGGCTGTTAAGTAT 2040
Qy 2563 GAGGGGAGAGAGATGAGAAATCTTTCAGTCCGCGCTGCTCAAGGGAGAACTG 2622
Db 2041 GAGGGGAGAGAGATGAGAAATCTTTCAGTCCGCGCTGCTCAAGGGAGAACTG 2100
Qy 2623 ATGCAATGGGTCCCGTCCCTGACAGTGTGCTGACAGCACAATCTGAGTCCCTCAC 2682
Db 2101 ATGCAATGGGTCCCGTCCCTGACAGTGTGCTGACAGCACAATCTGAGTCCCTCAC 2160
Qy 2683 TGCCTTGAAGAGACTGGGAAACAAAGCTTCTTCAACAAATGATGGCCATTGGAAGAA 2742
Db 2161 TGCCTTGAAGAGACTGGGAAACAAAGCTTCTTCAACAAATGATGGCCATTGGAAGAA 2220
Qy 2743 ATGGGATGATGATGAGAAACAGACATGGAGCTTATGATGCACTTTGCAATTAATTC 2802
Db 2221 ATGGGATGATGATGAGAAACAGACATGGAGCTTATGATGCACTTTGCAATTAATTC 2280
Qy 2803 AGCCGACACTGAGAGAGCTTTCAGTGAATGAGGGCAGGACAGATCAATGAGAGC 2862
Db 2281 AGCCGACACTGAGAGAGCTTTCAGTGAATGAGGGCAGGACAGATCAATGAGAGC 2340
Qy 2863 CCCACATGATGATCTGTTCAAGTGGGTCCAGTCAAGATGCTATTTGGCAGATTTCTC 2922
Db 2341 CCCACATGATGATCTGTTCAAGTGGGTCCAGTCAAGATGCTATTTGGCAGATTTCTC 2400
Qy 2923 TTCTCGCTCTCAAGGTCAACAGAAACCTGAAGAGTGAACCTAATGAGAACTGGCTG 2982
Db 2401 TTCTCGCTCTCAAGGTCAACAGAAACCTGAAGAGTGAACCTAATGAGAACTGGCTG 2460
Qy 2983 AGCCACTCTGACATGAGAGTCTTGTGAAGACCTGAGAGCGCCCTGCTGCTGAGAG 3042
Db 2461 AGCCACTCTGACATGAGAGTCTTGTGAAGACCTGAGAGCGCCCTGCTGCTGAGAG 2520
Qy 3043 ACCCTGCGTGGTGGCTGCTGCTGCTCAAGCTGAGAGCTGCAAGACCTTTCCTTTGGG 3102
Db 2521 ACCCTGCGTGGTGGCTGCTGCTGCTCAAGCTGAGAGCTGCAAGACCTTTCCTTTGGG 2580
Qy 3103 CTGAAGACCAACAGACCTGAGACCTGAGACCTGAGCTTCAATGAGCTCAAGAGTCT 3162
Db 2581 CTGAAGACCAACAGACCTGAGACCTGAGACCTGAGCTTCAATGAGCTCAAGAGTCT 2640
Qy 3163 GAGAGCAACACCTTTTCCAGAGACTGAGACAGCCGAGCTGCAAGCTCAACAGACTGAG 3222
Db 2641 GAGAGCAACACCTTTTCCAGAGACTGAGACAGCCGAGCTGCAAGCTCAACAGACTGAG 2700
Qy 3223 CTGCTGACGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
Db 2701 CTGCTGACGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Qy 3283 AGCCGACCTGAAGAGAGCTGAGACCTGAGAGAAACCTGAGATGAGCTTGGGCTGAG 3342
Db 2761 AGCCGACCTGAAGAGAGCTGAGACCTGAGAGAAACCTGAGATGAGCTTGGGCTGAG 2820
Qy 3343 CTGCTGATGAGAGGCTCAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3402
Db 2821 CTGCTGATGAGAGGCTCAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880

Qy 3403 ACACCTCTGAGATGAGATGAGGCAAGAACTGAGGGCCCTTGGAGCAGAGAACTCTGAG 3462
Db 2881 ACACCTCTGAGATGAGATGAGGCAAGAACTGAGGGCCCTTGGAGCAGAGAACTCTGAG 2940
Qy 3463 CTGCTCATCTTCAAGCAGAGCAAGAAACCAAGTGTGATGACCCCTACTGAGGGCCCTGATAG 3522
Db 2941 CTGCTCATCTTCAAGCAGAGCAAGAAACCAAGTGTGATGACCCCTACTGAGGGCCCTGATAG 3000
Qy 3523 GAGAGATGAGATGATGACATCTCTCACTCAAGCCGACAGAGACTGCGATCAGAGAGGGG 3582
Db 3001 GAGAGATGAGATGATGACATCTCTCACTCAAGCCGACAGAGACTGCGATCAGAGAGGGG 3060
Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGAGCTGAGCAAGATCTTCCCAATT 3642
Db 3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGAGCTGAGCAAGATCTTCCCAATT 3120
Qy 3643 GCTGAGATGAGAGAAAGCTCCCAAGAGATGATACCGGTGAAACTCTTGTGCGGCT 3702
Db 3121 GCTGAGATGAGAGAAAGCTCCCAAGAGATGATACCGGTGAAACTCTTGTGCGGCT 3180
Qy 3703 TCTCTGCTCTCAAGGGACCTGATACAGAGCTTTGGGACTGACATGACTTCTG 3762
Db 3181 TCTCTGCTCTCAAGGGACCTGATACAGAGCTTTGGGACTGACATGACTTCTG 3240
Qy 3763 GGGCCCAAGGGGCTGTGGCTACTGAGATGATGACAAAGAAAGAACTTTGACAGATT 3822
Db 3241 GGGCCCAAGGGGCTGTGGCTACTGAGATGATGACAAAGAAAGAACTTTGACAGATT 3300
Qy 3823 CACTTCCCTGATGAGGCTCTTACCGCTGGCCCAACAGGGGCTCTGCTTGTGATGAGA 3882
Db 3301 CACTTCCCTGATGAGGCTCTTACCGCTGGCCCAACAGGGGCTCTGCTTGTGATGAGA 3360
Qy 3883 GAAGCGGTACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3942
Db 3361 GAAGCGGTACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
Qy 3943 CCACAGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4002
Db 3421 CCACAGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
Qy 4003 GTGGAAGCTGTGACCTCCCTCACTTGTGAGCTTCCAGAGGGGCCATGTGACATCC 4062
Db 3481 GTGGAAGCTGTGACCTCCCTCACTTGTGAGCTTCCAGAGGGGCCATGTGACATCC 3540
Qy 4063 CTGTTCAAAATGAGCCCACTTTAAAGAGAGGGATGCTCTGAGAGAGCCAGCAGGGTG 4122
Db 3541 CTGTTCAAAATGAGCCCACTTTAAAGAGAGGGATGCTCTGAGAGAGCCAGCAGGGTG 3600
Qy 4123 GAGCTGATCAATGATTTTGAAGAAACCCAGCTTCTCCCTTGGGAGTCCCTCTGAAA 4182
Db 3601 GAGCTGATCAATGATTTTGAAGAAACCCAGCTTCTCCCTTGGGAGTCCCTCTGAAA 3660
Qy 4183 ATGATCATATGAGCCCTGAGCTTCAATCCGTCACCTGCTGATGATGATGATGATGATGAT 4242
Db 3661 ATGATCATATGAGCCCTGAGCTTCAATCCGTCACCTGCTGATGATGATGATGATGATGAT 3720
Qy 4243 GTTCATCTGAGAGAGTCACTTCACTCTTACCTGATCCAGATGACTGCTCATTTGG 4302
Db 3721 GTTCATCTGAGAGAGTCACTTCACTCTTACCTGATCCAGATGACTGCTCATTTGG 3780
Qy 4303 AAGGAACTGAGAGTCTGCTATGAGAGCCCTGAGAGAAACCAAGCTGTTCTCGAGTCTTAC 4362
Db 3781 AAGGAACTGAGAGTCTGCTATGAGAGCCCTGAGAGAAACCAAGCTGTTCTCGAGTCTTAC 3840
Qy 4363 GTTGGCACTTGGGATCAGGAGATCAGGCTGCAAGTGAAGAACAGAAAGATGAGACTCTG 4422
Db 3841 GTTGGCACTTGGGATCAGGAGATCAGGCTGCAAGTGAAGAACAGAAAGATGAGACTCTG 3900
Qy 4423 GTTGGAGAGGCTTGTGTAACAGAGAGATCTTATGCTTGAATCTGATCTCTCA 4482
Db 3901 GTTGGAGAGGCTTGTGTAACAGAGAGATCTTATGCTTGAATCTGATCTCTCA 3960

QY 4483 GCCCGCATAGCCGTAACCTTCACTTGGATGCCCCGAGTTGTCGACTTTGTGGAACAG 4542
DB 3961 GCCCGCATAGCCGTAACCTTCACTTGGATGCCCCGAGTTGTCGACTTTGTGGAACAG 4020
QY 4543 TATCGAGAGCAGTGTATAGCCCGATGATCATCGGTGAGGTTGTCTTGGACAACTGAT 4602
DB 4021 TATGAGAGAGCAGTGTATAGCCCGATGATCATCGGTGAGGTTGTCTTGGACAACTGAT 4080
QY 4603 GGACAGGTGCTGAGCCAGAGCAGTACAGAGGGTGTGCTGAGAACAGAGGCCAGC 4662
DB 4081 GGACAGGTGCTGAGCCAGAGCAGTACAGAGGGTGTGCTGAGAACAGAGGCCAGC 4140
QY 4663 CAGATGAGGAAAGCTGTTCAAGTTAGCCAGTCCGGAGCCGGAGTGAAGATGAGCTC 4722
DB 4141 CAGATGAGGAAAGCTGTTCAAGTTAGCCAGTCCGGAGCCGGAGTGAAGATGAGCTC 4200
QY 4723 TACCAAGCCCTGAGAGAGACCATCTCACTTATAGAACTCTGAGAGAGGAGC 4782
DB 4201 TACCAAGCCCTGAGAGAGACCATCTCACTTATAGAACTCTGAGAGAGGAGC 4260
QY 4783 AAAAAGGAACTCTGCACTCAGCAGC 4809
DB 4261 AAAAAGGAACTCTGCACTCAGCAGC 4287

RESULT 7

US-09-931-071-5
; Sequence 5, Application US/099331071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emed S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 0734-335001
; CURRENT APPLICATION NUMBER: US/09/931.071
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-931-071-5

Query Match 78.7%; Score 4287; DB 10; Length 4287;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 ATGCTGCGGAGCCTGGGCGCGCTGAGCTGTTACTTGAAGTTCTTGAAGAGAGAG 582
DB 1 ATGCTGCGGAGCCTGGGCGCGCTGAGCTGTTACTTGAAGTTCTTGAAGAGAGAG 60
QY 583 CTGAAGAGAGTTCAGCTTCTGCTGCGCAATAAGAGCACTCCAGAGAGTCTTGGGTAG 642
DB 61 CTGAAGAGAGTTCAGCTTCTGCTGCGCAATAAGAGCACTCCAGAGAGTCTTGGGTAG 120
QY 643 ACACCCGCTGACGAGAGAGAGAGAGAGTGGATGAGAGTGGCTGTAACCTGATGAGTCTG 702
DB 121 ACACCCGCTGACGAGAGAGAGAGAGTGGATGAGAGTGGCTGTAACCTGATGAGTCTG 180
QY 703 TATGGGAGAGCAGCGGCTGAGAGCTTACCTTCAATCTGAGAGCAGATGAGGCTGAG 762
DB 181 TATGGGAGAGCAGCGGCTGAGAGCTTACCTTCAATCTGAGAGCAGATGAGGCTGAG 240
QY 763 TCACTGTGCGCCAGAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 822
DB 241 TCACTGTGCGCCAGAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCAACCACTCCAGCGAGTCTATGCGCTGG 882

DB 301 AGTGAACCCCACTGGGGTCTCCAGCAACCACTCCAGCGAGTCTATGCGCTGG 360
QY 883 ATCCATGAATTTGCCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGAACAGCTG 942
DB 361 ATCCATGAATTTGCCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGAACAGCTG 420
QY 943 CTGACACATCTGAGAGCGCTGAGAGAAATCTCTGCTCACTCTTCACTCAAGCTCTT 1002
DB 421 CTGACACATCTGAGAGCGCTGAGAGAAATCTCTGCTCACTCTTCACTCAAGCTCTT 480
QY 1003 CCAAGCTCCCAAGACATGAGTCTCAAGCCAGAGTCAACCAAGCCGCCACATCCACA 1062
DB 481 CCAAGCTCCCAAGACATGAGTCTCAAGCCAGAGTCAACCAAGCCGCCACATCCACA 540
QY 1063 GCACTGCTGGGAGAGCTGGGATCCCACTCAGCCAGCTTACGACCCAGAGAGAG 1122
DB 541 GCACTGCTGGGAGAGCTGGGATCCCACTCAGCCAGCTTACGACCCAGAGAGAG 600
QY 1123 GCTCTGGGAGCCCAATGGGCTCTGAGATGAACAGTCAAGAAATTACTACAGAAATCAGA 1182
DB 601 GCTCTGGGAGCCCAATGGGCTCTGAGATGAACAGTCAAGAAATTACTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGAGAGGCCCCATGGGAGGCTGTGAGAACG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGAGAGGCCCCATGGGAGGCTGTGAGAACG 720
QY 1243 CCCCCACAGGCGCACACAGCTTACAGCCCAACCAACCACTGAGAGGCTTCTGTAGA 1302
DB 721 CCCCCACAGGCGCACACAGCTTACAGCCCAACCAACCACTGAGAGGCTTCTGTAGA 780
QY 1303 GAGAGGCTCTGTTCCATGAGGCTCTGAGAAATGAGAAATTTAACAATAATTCACAG 1362
DB 781 GAGAGGCTCTGTTCCATGAGGCTCTGAGAAATGAGAAATTTAACAATAATTCACAG 840
QY 1363 CTGCTACTTCTACAAAGACCTCAACCCAGAGAGCAAGTCCCTGTCTCAAGAGAGCTG 1422
DB 841 CTGCTACTTCTACAAAGACCTCAACCCAGAGAGCAAGTCCCTGTCTCAAGAGAGCTG 900
QY 1423 CTGATTAATGTGAGAGAAATCAGAGCAATTAATGAGATCAGAGCTTATTTGGCCCA 1482
DB 901 CTGATTAATGTGAGAGAAATCAGAGCAATTAATGAGATCAGAGCTTATTTGGCCCA 960
QY 1483 GGCTGTGATCCCAAGAACCTGCGATGATGATCTGAGAGGGGGTGTGATGAGAG 1542
DB 961 GGCTGTGATCCCAAGAACCTGCGATGATGATCTGAGAGGGGGTGTGATGAGAG 1020
QY 1543 TCAACACTGGCCAGGAGAGTGAAGAGAGCTGGGGAGAGGCGCATGTATGGGAGCCGC 1602
DB 1021 TCAACACTGGCCAGGAGAGTGAAGAGAGCTGGGGAGAGGCGCATGTATGGGAGCCGC 1080
QY 1603 TTCCAGCATGTTCTTCTACTTCAAGCTGCAAGAGAGCTGGCCCATGAGAGTGTGATCTC 1662
DB 1081 TTCCAGCATGTTCTTCTACTTCAAGCTGCAAGAGAGCTGGCCCATGAGAGTGTGATCTC 1140
QY 1663 GCTGAGCTCATGAGAAAGATGGGAGAGCACTCGGCTCCCATTTAGACAGATCTGTCT 1722
DB 1141 GCTGAGCTCATGAGAAAGATGGGAGAGCACTCGGCTCCCATTTAGACAGATCTGTCT 1200
QY 1723 AGGCGAGAGCGCTGCTTCTCATCTCGATGTGTATGATGAGCCAGATGAGGCTTTCAG 1782
DB 1201 AGGCGAGAGCGGCTGCTTCTCATCTCGATGTGTATGATGAGCCAGATGAGGCTTTCAG 1260
QY 1783 GAGCGAGATGTGAGCTGTGTCTGCACTGAGAGCAGAGCAGCGGGGAGATGCACTGCTG 1842
DB 1261 GAGCGAGATGTGAGCTGTGTCTGCACTGAGAGCAGAGCAGCGGGGAGATGCACTGCTG 1320
QY 1843 GGCAGTTTGTGGGAGAAACATATCTTCCGAGGAGATCTTCTGATCAAGGCTCGAGC 1902
DB 1321 GGCAGTTTGTGGGAGAAACATATCTTCCGAGGAGATCTTCTGATCAAGGCTCGAGC 1380
QY 1903 ACACTCTGAGAACTCATATCTTCTTGTGAGAGAGCAGCTTGGGTGAGGCTCTGGGG 1962
DB 1381 ACACTCTGAGAACTCATATCTTCTTGTGAGAGAGCAGCTTGGGTGAGGCTCTGGGG 1440

QY 1963 TTCTGAGTCCAGCAGAGAAATATTTCTACAGATTTTCAAGATGAAAGCAAGCA 2022
 DB 1441 TTCTGAGTCCAGCAGAGAAATATTTCTACAGATTTTCAAGATGAAAGCAAGCA 1500
 QY 2023 ATTAGAGCCTTTAGTGTGATCAATCAACAAAGAGCTGTGGCCCTGTGCTTGTGCC 2082
 DB 1501 ATTAGAGCCTTTAGTGTGATCAATCAACAAAGAGCTGTGGCCCTGTGCTTGTGCC 1560
 QY 2083 TGGGTGTCTGGCTGGCTGCTCACTTGGCTGATGAGCAGATGAAAGCGAAAGAAATCTC 2142
 DB 1561 TGGGTGTCTGGCTGGCTGCTCACTTGGCTGATGAGCAGATGAAAGCGAAAGAAATCTC 1520
 QY 2143 ACACTGATCTTCCAAAGACCAACCAACCTCTGTCTATTAATTAATTAATTAATTAATTA 2202
 DB 1621 ACACTGATCTTCCAAAGACCAACCAACCTCTGTCTATTAATTAATTAATTAATTAATTA 1680
 QY 2203 GCTCAGCATTGGGAGACCCAGCTCAGAGACCTGTCTGTGGCTGTGAGGAGATCTGG 2262
 DB 1681 GCTCAGCATTGGGAGACCCAGCTCAGAGACCTGTCTGTGGCTGTGAGGAGATCTGG 1740
 QY 2263 CAAAAAAGACCTTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2322
 DB 1741 CAAAAAAGACCTTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 QY 2323 ATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTTACAGC 2382
 DB 1801 ATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTTACAGC 1860
 QY 2383 TTCAATCACCTCTGTTTCCAAAGATTTCTTGAAGACATGCTATGCTTGAAGATGAG 2442
 DB 1861 TTCAATCACCTCTGTTTCCAAAGATTTCTTGAAGACATGCTATGCTTGAAGATGAG 1920
 QY 2443 AAGGGAGAGATTAACATTTATTTGATCATATGATTTGAAAAAGACGTAGAAGCATAT 2502
 DB 1921 AAGGGAGAGATTAACATTTATTTGATCATATGATTTGAAAAAGACGTAGAAGCATAT 1980
 QY 2503 GGAATACATGGCTCTGTTTGGGGATCAACACAGTTTCTTATTTGGGCTCTTAAAGAT 2562
 DB 1981 GGAATACATGGCTCTGTTTGGGGATCAACACAGTTTCTTATTTGGGCTCTTAAAGAT 2040
 QY 2563 GAGGGGAGAGAGATGAGAAATCTTCACTGCGGCTGTCTCAGGGAGAAAGCTG 2622
 DB 2041 GAGGGGAGAGAGATGAGAAATCTTCACTGCGGCTGTCTCAGGGAGAAAGCTG 2100
 QY 2623 ATGCAATGGGTCCCGTCCCTGAGCTGCTGAGCAGACATCTCTGAGATCTCTCAC 2682
 DB 2101 ATGCAATGGGTCCCGTCCCTGAGCTGCTGAGCAGACATCTCTGAGATCTCTCAC 2160
 QY 2683 TGGCTTGAAGAGACTCGGAAACAAAGCTTCTGACACAAAGTATGGGCCATTTTGAAGAA 2742
 DB 2161 TGGCTTGAAGAGACTCGGAAACAAAGCTTCTGACACAAAGTATGGGCCATTTTGAAGAA 2220
 QY 2743 ATGGGACATGTGTGAGAAACAGACATGAGACTCTTAAGTGTGCACTTTTGAATTTTC 2280
 DB 2221 ATGGGACATGTGTGAGAAACAGACATGAGACTCTTAAGTGTGCACTTTTGAATTTTC 2280
 QY 2803 AGCGCCACGTGAAGAACTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGAGAGC 2862
 DB 2281 AGCGCCACGTGAAGAACTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGAGAGC 2340
 QY 2863 CCAACCATGATATCTCTGTTCAAGTGGGTCCCACTGACAGATGCTTATTTGGCAGATTTCTC 2922
 DB 2341 CCAACCATGATATCTCTGTTCAAGTGGGTCCCACTGACAGATGCTTATTTGGCAGATTTCTC 2400
 QY 2923 TTCTCGCTCTCAAGGTCAACAGAACTGAAAGAGCTGAGACCTTAAGTGAATCTGCTG 2982
 DB 2401 TTCTCGCTCTCAAGGTCAACAGAACTGAAAGAGCTGAGACCTTAAGTGAATCTGCTG 2460
 QY 2983 AGCCACTCTGCAAGTAAAGATCTTTGTAAGACCTTGAAGAGCTCTGCTCTCTGAGAG 3042
 DB 2461 AGCCACTCTGCAAGTAAAGATCTTTGTAAGACCTTGAAGAGCTCTGCTCTCTGAGAG 2520

QY 3043 ACCCTGGGTGGCTGGCTGGCTGCTCAGAGCTGAGAGCTGCAAGGACCTTGGCTTGGG 3102
 DB 2521 ACCCTGGGTGGCTGGCTGGCTGCTCAGAGCTGAGAGCTGCAAGGACCTTGGCTTGGG 2580
 QY 3103 CTGAGAGCCAAACAGACCTGACCGAGCTGAGACTGAGCTTCAATGTGCTCAGAGATGCT 3162
 DB 2581 CTGAGAGCCAAACAGACCTGACCGAGCTGAGACTGAGCTTCAATGTGCTCAGAGATGCT 2640
 QY 3163 GAGGCCAAACACTTTTGCAGAGAGCTGAGACAGCCGAGCTGCAAGCTTACAGCACTGAG 3222
 DB 2641 GAGGCCAAACACTTTTGCAGAGAGCTGAGACAGCCGAGCTGCAAGCTTACAGCACTGAG 2700
 QY 3223 CTGGTACGCTGGGCTGCTCAGTCTGATCTGCTGCAAGAGCCTGGCTCTGTGCTTAAGTCC 3282
 DB 2701 CTGGTACGCTGGGCTGCTCAGTCTGATCTGCTGCAAGAGCCTGGCTCTGTGCTTAAGTCC 2760
 QY 3283 AGCCCAAGCTGAGAGAGCTAGACTGAGCAGAGAAACCTGATGACGTTGGCTGCGA 3342
 DB 2761 AGCCCAAGCTGAGAGAGCTAGACTGAGCAGAGAAACCTGATGACGTTGGCTGCGA 2820
 QY 3343 CTGCTGTGAGGGGCTCAGGACATCTGCTGCAAACTCATACGCTGGGGCTGAGACAG 3402
 DB 2821 CTGCTGTGAGGGGCTCAGGACATCTGCTGCAAACTCATACGCTGGGGCTGAGACAG 2880
 QY 3403 ACAACTGTAGATGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3462
 DB 2881 ACAACTGTAGATGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
 QY 3463 CTGCTCATCTTCCAGCAG 3522
 DB 2941 CTGCTCATCTTCCAGCAG 3000
 QY 3523 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3582
 DB 3001 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
 QY 3583 GCTTCCATGTTGCTCAGGCTATCTCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 3642
 DB 3061 GCTTCCATGTTGCTCAGGCTATCTCAAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 3120
 QY 3643 GCTGAGATTTGAG 3702
 DB 3121 GCTGAGATTTGAG 3180
 QY 3703 TCTCTGCTCTCAG 3762
 DB 3181 TCTCTGCTCTCAG 3240
 QY 3763 GGGCCCAAGGGGCTGTGGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3822
 DB 3241 GGGCCCAAGGGGCTGTGGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
 QY 3823 CACTTCCCTGATGAGCTGCTCTTACAGCTGAGCCCAACAGCGGCTCTCTTGTGATGAG 3882
 DB 3301 CACTTCCCTGATGAGCTGCTCTTACAGCTGAGCCCAACAGCGGCTCTCTTGTGATGAG 3360
 QY 3883 GAAAGCGTGAACGTTGAGATTTGAATTTCTGTGTGAGAGACAGTTCTGGGTGAGATCAAC 3942
 DB 3361 GAAAGCGTGAACGTTGAGATTTGAATTTCTGTGTGAGAGACAGTTCTGGGTGAGATCAAC 3420
 QY 3943 CCAACAGCAGAGCTGATGAG 4002
 DB 3421 CCAACAGCAGAGCTGATGAG 3480
 QY 4003 GTGGAAGCTGTGACCTTCTCTCACTTGTGTGCTTCAAGGGGAGATGAGACATCC 4062
 DB 3481 GTGGAAGCTGTGACCTTCTCTCACTTGTGTGCTTCAAGGGGAGATGAGACATCC 3540
 QY 4063 CTGTTCCTTCAATGGCCCACTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4122
 DB 3541 CTGTTCCTTCAATGGCCCACTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
 QY 4123 GAGCTCATCACTAGATTTGAGAAACCCAGCTTCTCCCCCTTGGAGATCTCTGAAA 4182

Db 3601 GAGCTGATCATCATAGTTCTGGAAAACCCAGCTTCTCCCTTGGAGTCTCTGAAA 3660
Qy 4183 ATGATTCATATATGCTCCCTGCGCTTCAATTCCTGACCTCTGTGTGTGTTTACACCGC 4242
Db 3661 ATGATTCATATATGCTCCCTGCGCTTCAATTCCTGACCTCTGTGTGTGTTTACACCGC 3720
Qy 4243 GTCCATCTGAGGAAGTCACTTCCACTTCACTGATCCCAAGTGACTGCTCATTCGG 4302
Db 3721 GTCTATCTGAGGAAGTCACTTCCACTTCACTGATCCCAAGTGACTGCTCATTCGG 3780
Qy 4303 AAGGAATCTGAGCTCTGTATGAGACCCCTGAGAGACCACTGTTCTCGAGTTCTAC 4362
Db 3781 AAGGAATCTGAGCTCTGTATGAGACCCCTGAGAGACCACTGTTCTCGAGTTCTAC 3840
Qy 4363 GTTGGCCACTTGGAGATCAGGGATCAGGCTGCAATGTAAGCAAGAAAGATGAGACTCTG 4422
Db 3841 GTTGGCCACTTGGAGATCAGGGATCAGGCTGCAATGTAAGCAAGAAAGATGAGACTCTG 3900
Qy 4423 GTGTGGAGGCTTGGTGAACAGAGATCTCATGCTGCAACTGATCCCTGCA 4482
Db 3901 GTGTGGAGGCTTGGTGAACAGAGATCTCATGCTGCAACTGATCCCTGCA 3960
Qy 4483 GCCCGCATAGCCCTTACCTTCACTCTGAGTCCCGCAGTGTGCACTTTGTGAGCCAG 4542
Db 3961 GCCCGCATAGCCCTTACCTTCACTCTGAGTCCCGCAGTGTGCACTTTGTGAGCCAG 4020
Qy 4543 TATGAGAGAGCTGATGAGCCCGAGTGAATCGGTGAGGTTGTTTGAACAACTGCAT 4602
Db 4021 TATGAGAGAGCTGATGAGCCCGAGTGAATCGGTGAGGTTGTTTGAACAACTGCAT 4080
Qy 4603 GGACAGGCTGAGCCAGGAGACAGTGAAGAGGCTGCTGAGAACAGAGGCCCAAGC 4662
Db 4081 GGACAGGCTGAGCCAGGAGACAGTGAAGAGGCTGCTGAGAACAGAGGCCCAAGC 4140
Qy 4663 CAGATGCGGAGAGCTGTTCAAGCTTGAGCAGTCTCTGAGACCGGAAGTGAAGATGACTC 4722
Db 4141 CAGATGCGGAGAGCTGTTCAAGCTTGAGCAGTCTCTGAGACCGGAAGTGAAGATGACTC 4200
Qy 4723 TACCAAGCTTGAAGAGAGACCACTTCACTTATGAACTCTGAGAGAGAGGCAAGC 4782
Db 4201 TACCAAGCTTGAAGAGAGACCACTTCACTTATGAACTCTGAGAGAGAGGCAAGC 4260
Qy 4783 AAAAAGGAGCTCTGCGCACTGAGAGC 4809
Db 4261 AAAAAGGAGCTCTGCGCACTGAGAGC 4287

RESULT 8

US-09-388-221-1

; Sequence 1, Application US/09388221A

; Publication No. US20020192643A1

; GENERAL INFORMATION:

; APPLICANT: Reed, John C.

; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regul

; FILE REFERENCE: P-1f 3650

; CURRENT APPLICATION NUMBER: US/09/388,221A

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 4422

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(4422)

US-09-388-221-1

Query Match 76.2%; Score 4148; DB 10; Length 4422;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

Qy 523 ATGAGTGGCGAGAGCTGGGCGCGCTGCTGTTACTTGGAGTTCTGAAGAGAGAG 582
Db 1 ATGAGTGGCGAGAGCTGGGCGCGCTGCTGTTACTTGGAGTTCTGAAGAGAGAGAG 60
Qy 583 CTGAAGAGATTCCAGCTTCTGTGCGCAATPAAGCGCACTCCAGAGCTTTCGGGTGAG 642
Db 61 CTGAAGAGATTCCAGCTTCTGTGCGCAATPAAGCGCACTCCAGAGCTTTCGGGTGAG 120
Qy 643 ACACCCGCTCAGCCAGAGAGACGAGTGGATGAGAGGTGGCTCTGTACTGGTGGCTCAG 702
Db 121 ACACCCGCTCAGCCAGAGAGACGAGTGGATGAGAGGTGGCTCTGTACTGGTGGCTCAG 180
Qy 703 TATGGAGAGAGCGGCGCTGGAGCTTACCTCATCTCTGAGAGAGATGGGCTGAG 762
Db 181 TATGGAGAGAGCGGCGCTGGAGCTTACCTCATCTCTGAGAGAGATGGGCTGAG 240
Qy 763 TCACTGTGCGCCAGAGCCAGAGAGGGGAGGCCACTCTCTCTCATTTCCCTACAGCCCA 822
Db 241 TCACTGTGCGCCAGAGCCAGAGAGGGGAGGCCACTCTCTCTCATTTCCCTACAGCCCA 300
Qy 823 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGGAGTGTAAATGCCCTGG 882
Db 301 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGGAGTGTAAATGCCCTGG 360
Qy 883 ATCCATGATTTGCGGCGGGGTGCAACCAAGGCTCAGAGAGAGGTTTGAAGACAGCTG 942
Db 361 ATCCATGATTTGCGGCGGGGTGCAACCAAGGCTCAGAGAGAGGTTTGAAGACAGCTG 420
Qy 943 CTGAGCACTATGAGCGCGCTGAGAGAAATCTGTGCTTCACTCTCTTCAACAGTCTTT 1002
Db 421 CTGAGCACTATGAGCGCGCTGAGAGAAATCTGTGCTTCACTCTCTTCAACAGTCTTT 480
Qy 1003 CCAAGCTTCCCAAGCATGAGTCTTCAAGCCAGAGATCAACCAAGCCGCCCATCTCCACA 1062
Db 481 CCAAGCTTCCCAAGCATGAGTCTTCAAGCCAGAGATCAACCAAGCCGCCCATCTCCACA 540
Qy 1063 GCAGTGCTGGGAGCTGGGATCCCACTCAGGCCAGCTTGAAGCCAGAGACAGAG 1122
Db 541 GCAGTGCTGGGAGCTGGGATCCCACTCAGGCCAGCTTGAAGCCAGAGACAGAG 600
Qy 1123 GCTCTTGGGAGCCCAATGAGCTTGTGATGAACGTGAGAAATTTTACTACAGAAATCAGA 1182
Db 601 GCTCTTGGGAGCCCAATGAGCTTGTGATGAACGTGAGAAATTTTACTACAGAAATCAGA 660
Qy 1183 GAAGAGAGAGAGAAATCAGAGAAAGCAGAGCCCAATGGGAGCGGTGTGGAAGC 1242
Db 661 GAAGAGAGAGAGAAATCAGAGAAAGCAGAGCCCAATGGGAGCGGTGTGGAAGC 720
Qy 1243 CCCCCAGGCGGACACAGAGCTTACAGGCCCAACCAAGCCATGGAGCTTCTGTGAGA 1302
Db 721 CCCCCAGGCGGACACAGAGCTTACAGGCCCAACCAAGCCATGGAGCTTCTGTGAGA 780
Qy 1303 GAGAGCTTGTTCACATGGCCCTGGAATAATGAGATTTTAAACAAAATTCACACAG 1362
Db 781 GAGAGCTTGTTCACATGGCCCTGGAATAATGAGATTTTAAACAAAATTCACACAG 840
Qy 1363 CTGCTACTTTCACAAAGCTTCAACCCAGAGCCAAATCCCCGTGTCAAGAGAGCTGG 1422
Db 841 CTGCTACTTTCACAAAGCTTCAACCCAGAGCCAAAGTCCCTGTGTCAAGAGAGCTGG 900
Qy 1423 CCTGATTAATGAGAGAGAAATCAGAGCACTTAATGAGATCAGAGCTTATTTGGCCCA 1482
Db 901 CCTGATTAATGAGAGAGAAATCAGAGCACTTAATGAGATCAGAGCTTATTTGGCCCA 960
Qy 1483 GGCCTGATTAACCAAGAACTTGATAGTCACTGAGAGGGGCTGTGTAATTTGGGAAG 1542
Db 961 GGCCTGATTAACCAAGAACTTGATAGTCACTGAGAGGGGCTGTGTAATTTGGGAAG 1020
Qy 1543 TCAACACTGGCCAGAGAGTGAAGAAAGCCCTGGGGAGAGGCCAGCTGTATGGGGACCGC 1602
Db 1021 TCAACACTGGCCAGAGAGTGAAGAAAGCCCTGGGGAGAGGCCAGCTGTATGGGGACCGC 1080
Qy 1603 TTCAGAGATGCTTCTACTTCACTGAGCTGAGAGAGCTGGCCCAAGTGTGTGAGTCTC 1662

Qy 3823 CACTTCCCTAGTGGCTCTTACCGCTGGCCAAACAGGGGCTCTGCTTTGTATGAGA 3882
 Db 3301 CACTTCCCTAGTGGCTCTTACCGCTGGCCAAACAGGGGCTCTGCTTTGTATGAGA 3360
 Qy 3883 GAAGCGGTGACCGTTGAGATTGAATTCGTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC 3942
 Db 3361 GAAGCGGTGACCGTTGAGATTGAATTCGTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC 3420
 Qy 3943 CCACAGCAGAGCTGATGTGTGGCAGGGGCTCTGCTGACATCAAGAGCTGAGCTGAGACT 4002
 Db 3421 CCACAGCAGAGCTGATGTGTGGCAGGGGCTCTGCTGACATCAAGAGCTGAGACTGAGACT 3480
 Qy 4003 GTGGAAGCTGAGCCTCCCTCACTTGTGGCTCCCAAGGGGGGCAATGTGAGACATCC 4062
 Db 3481 GTGGAAGCTGAGCCTCCCTCACTTGTGGCTCCCAAGGGGGGCAATGTGAGACATCC 3540
 Qy 4063 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGAGTCTCTGAGAAAGCCAGCGAGTG 4122
 Db 3541 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGAGTCTCTGAGAAAGCCAGCGAGTG 3600
 Qy 4123 GAGCTGCTCAATATGTTCTGAGAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAA 4182
 Db 3601 GAGCTGCTCAATATGTTCTGAGAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAA 3660
 Qy 4183 ATGATCCATATGCGCTGCGCTTCACTCCGTCACCTGAGGTGTTGCTTACCAACCGC 4242
 Db 3661 ATGATCCATATGCGCTGCGCTTCACTCCGTCACCTGAGGTGTTGCTTACCAACCGC 3720
 Qy 4243 GTCCATCTGAGAGAGTCACTTCCACTCTTACCTGATCCCAAGTGAATGCTGCTCAATGG 4302
 Db 3721 GTCCATCTGAGAGAGTCACTTCCACTCTTACCTGATCCCAAGTGAATGCTGCTCAATGG 3780
 Qy 4303 ----- 4302
 Db 3781 AAGCCATAGATGATCTAGAATAATTCAGATTGTGCGAATCCAGAACCCACCCCG 3840
 Qy 4303 ----- 4302
 Db 3841 CTGACCCCATTTATATGAGGCTGTGCTTAACTGTGTCTGGGTCTGTTCAAGGATGCTG 3900
 Qy 4303 -----AAGAACTGAGGCTGTGCTTAACTGTGTCTGGGTCTGTTCAAGGATGCTG 4350
 Db 3901 GAATATCTCCCAAGAGATGAGAGCTGTGCTATGGAAGCCCTGAGAAAGACAGCTGTTC 3960
 Qy 4351 TCGAGATTCTACGTTGGCCACTTGGAGATCAGGAGATCAGGCTGCAAGTGAAGACAAAGAA 4410
 Db 3961 TCGAGATTCTACGTTGGCCACTTGGAGATCAGGAGATCAGGCTGCAAGTGAAGACAAAGAA 4020
 Qy 4411 GATGAGATCTGTGTGTGGAGGCTTGTGTGAAACCAAGAGATCTCATGCTCTGCAATCTA 4470
 Db 4021 GATGAGATCTGTGTGTGGAGGCTTGTGTGAAACCAAGAGATCTCATGCTCTGCAATCTA 4080
 Qy 4471 CTGATCCCTCAGAGCCGCAATAGCCGTAACCTCTGAGATGCCCCGCAATTTCTGAC 4530
 Db 4081 CTGATCCCTCAGAGCCGCAATAGCCGTAACCTCTGAGATGCCCCGCAATTTCTGAC 4140
 Qy 4531 TTTGTGACCAAGTATCGAGAGCAAGTATAGCCGAGTGAATCATGCTGTGAGAGTTGTGG 4590
 Db 4141 TTTGTGACCAAGTATCGAGAGCAAGTATAGCCGAGTGAATCATGCTGTGAGAGTTGTGG 4200
 Qy 4591 GACAAATCGATGACAGGTCTGAGCCAGAGAGATGACAGAGAGGTCTGTGAGAAC 4650
 Db 4201 GACAAATCGATGACAGGTCTGAGCCAGAGAGATGACAGAGAGGTCTGTGAGAAC 4260
 Qy 4651 ACGAGGCCCAAGCAGATGCGAGAGCTGTTCACTTGAAGCCAGTCTCTGAGAAC 4710
 Db 4261 ACGAGGCCCAAGCAGATGCGAGAGCTGTTCACTTGAAGCCAGTCTCTGAGAAC 4320
 Qy 4711 AAAGATGACTCTCAAAAGCCCTTGAAGAGAACCAATCTCTCACTTATATGAAACTCTGG 4770
 Db 4321 AAAGATGACTCTCAAAAGCCCTTGAAGAGAACCAATCTCTCACTTATATGAAACTCTGG 4380

Qy 4771 GAGAAGGCGAGCAAAAAGGAGCTCTGCTCACTCAGCAGCTGA 4812
 Db 4381 GAGAAGGCGAGCAAAAAGGAGCTCTGCTCACTCAGCAGCTGA 4422

 RESULT 9
 US-09-388-221-3
 ; Sequence 3, Application US/09388221A
 ; Publication No. US20020192643A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regul;
 ; FILE REFERENCE: P-LJ 3650
 ; CURRENT FILING DATE: 1999-09-01
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 4200
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(4197)
 US-09-388-221-3

 Query Match 75.3%; Score 4100; DB 10; Length 4200;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 4200; Conservative 0; Mismatches 0; Indels 90; Gaps 1;

 Qy 523 ATGAGCTGGAGGAGCTGGGGGCGGCTGGCGCTGCTGACTTGGAGTTCTGAAGAGAGAG 582
 Db 1 ATGAGCTGGAGGAGCTGGGGGCGGCTGGCGCTGCTGACTTGGAGTTCTGAAGAGAGAG 60
 Qy 583 CTGAAGAGTTTCCAGCTTGTGCTGCGCAATTAAGCGCACTCCAGAGCTTTGGGTGAG 642
 Db 61 CTGAAGAGTTTCCAGCTTGTGCTGCGCAATTAAGCGCACTCCAGAGCTTTGGGTGAG 120
 Qy 643 ACACCCGCTACGCCAGAGAGAGAGAGAGAGTGAAGTGGCTCTGTACTGTGTGGCTGAG 702
 Db 121 ACACCCGCTACGCCAGAGAGAGAGAGAGAGTGAAGTGGCTCTGTACTGTGTGGCTGAG 180
 Qy 703 TATGGGAGAGAGCGGGGCTGGGAGACTGAGCCCTCATCTCTGGAGAGATGGGCTGAG 762
 Db 181 TATGGGAGAGAGCGGGGCTGGGAGACTGAGCCCTCATCTCTGGAGAGATGGGCTGAG 240
 Qy 763 TCACTGTGCGCCAGGCCAGAGAGAGGAGGAGGCACTCTCCCTATTCCCTACAGGCCA 822
 Db 241 TCACTGTGCGCCAGGCCAGAGAGAGGAGGAGGCACTCTCCCTATTCCCTACAGGCCA 300
 Qy 823 AGTGAACCCCACTGGGAGTCTCCAGCCAAACCACTTCCAGCGAGTCTAATGCCCTGG 882
 Db 301 AGTGAACCCCACTGGGAGTCTCCAGCCAAACCACTTCCAGCGAGTCTAATGCCCTGG 360
 Qy 883 ATCCATGAATTGCGGGGAGGTGACCCAGAGGCTCAGAGAGAGGTTTGTGACAGCTG 942
 Db 361 ATCCATGAATTGCGGGGAGGTGACCCAGAGGCTCAGAGAGAGGTTTGTGACAGCTG 420
 Qy 943 CTTGACATCTGAGACGCGCTGAGAGAAATCTTCTGCTCACTCTTCAACAGCTTT 1002
 Db 421 CTTGACATCTGAGACGCGCTGAGAGAAATCTTCTGCTCACTCTTCAACAGCTTT 480
 Qy 1003 CCAAGCTTCCCAAGCACTAGTCTTCAAGCCAGAGATCAACCAAGCCCTCAATCCACA 1062
 Db 481 CCAAGCTTCCCAAGCACTAGTCTTCAAGCCAGAGATCAACCAAGCCCTCAATCCACA 540
 Qy 1063 GCAGTGTGGGAGAGCTGGGATCCCACTCAGGCCAGCTTGAACCCAGAGACAGAG 1122
 Db 541 GCAGTGTGGGAGAGCTGGGATCCCACTCAGGCCAGCTTGAACCCAGAGACAGAG 600
 Qy 1123 GCTCTGGGAGCCCAATGAGCTTGTGATGAAACGTCAAGAAATTTACTACAGAAATCAGA 1182
 Db 601 GCTCTGGGAGCCCAATGAGCTTGTGATGAAACGTCAAGAAATTTACTACAGAAATCAGA 660

QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGCGGCCCATGGGCGCTGTGGTAGAG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGGCGGCCCATGGGCGCTGTGGTAGAG 720
QY 1243 CCCCACAGGCGGACACAGCCTACAGCCCAACACCAACCAAGGAGCCCTTCGTAGA 1302
DB 721 CCCCACAGGCGGACACAGCCTACAGCCCAACCAACCAAGGAGCCCTTCGTAGA 780
QY 1303 GAGAGCCTCTGTTCACATGAGCCCTGGAAAAATGAGATTTTAAACAAAAATTCACAG 1362
DB 781 GAGAGCCTCTGTTCACATGAGCCCTGGAAAAATGAGATTTTAAACAAAAATTCACAG 840
QY 1363 CTGCTACTTCTACAAACCTCACCCCAAGGCCAATCCCTGGTCAAGAGAGCTGG 1422
DB 841 CTGCTACTTCTACAAACCTCACCCCAAGGCCAATCCCTGGTCAAGAGAGCTGG 900
QY 1423 CCGATTAATGAGAGAGAAATCGAGACATTTAATTGAGATCAGAGCTTAATTGGCCCA 1482
DB 901 CCGATTAATGAGAGAGAAATCGAGACATTTAATTGAGATCAGAGCTTAATTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACTCGCATAGTCATACAGAGGCGCTGCTGGAATTTGGAG 1542
DB 961 GGCCTGGATACCAAGAACTCGCATAGTCATACAGAGGCGCTGCTGGAATTTGGAG 1020
QY 1543 TCAACACTGGCCAGGCGAGGTGAGAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGACCG 1602
DB 1021 TCAACACTGGCCAGGCGAGGTGAGAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGACCG 1080
QY 1603 TTCAGACTGTCTTCTACTTCAAGCTGAGAGAGCTGGCCAGTCAGAGGTGAGAGCTC 1662
DB 1081 TTCAGACTGTCTTCTACTTCAAGCTGAGAGAGCTGGCCAGTCAGAGGTGAGAGCTC 1140
QY 1663 GCTGAGCTCATCGAAAAAGATGGAGACAGCACTCCGCTCCCATTTAGACAGATCTGTCT 1722
DB 1141 GCTGAGCTCATCGAAAAAGATGGAGACAGCACTCCGCTCCCATTTAGACAGATCTGTCT 1200
QY 1723 AGGCGTAAGGGGCTGCTCTTCTCTGATGGTGTATGATGACCAAGATGGGCTTTGAG 1782
DB 1201 AGGCGTAAGGGGCTGCTCTTCTCTGATGGTGTATGATGACCAAGATGGGCTTTGAG 1260
QY 1783 GAGCGAGGTTCTGAGCTGTGCTGCACTGAGCAGGACAGCCGCGGAGATGCACTGCTG 1842
DB 1261 GAGCGAGGTTCTGAGCTGTGCTGCACTGAGCAGGACAGCCGCGGAGATGCACTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAACTATATCTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1902
DB 1321 GGCAGTTTGTGGGAAAACTATATCTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1380
QY 1903 ACAGCTCTGAGAACTTCACTTCTTGTGAGAGGACAGTTGGGTGAGAGTCTTGGGG 1962
DB 1381 ACAGCTCTGAGAACTTCACTTCTTGTGAGAGGACAGTTGGGTGAGAGTCTTGGGG 1440
QY 1963 TTCTCTAGTCTCAGCAGAGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 2022
DB 1441 TTCTCTAGTCTCAGCAGAGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 1500
QY 2023 ATTAGACCTTTTAGTGTGTCAATCAAAACAAGAGCTTGGGCGCTGTGTCTTGTGCC 2082
DB 1501 ATTAGACCTTTTAGTGTGTGTCAATCAAAACAAGAGCTTGGGCGCTGTGTGTGCC 1560
QY 2083 TGGGTGTCTGGGCTGCTGCACTTGGCTGATGAGAGATGAAAGCGGAAAGAAAACTC 2142
DB 1561 TGGGTGTCTGGGCTGCTGCACTTGGCTGATGAGAGATGAAAGCGGAAAGAAAACTC 1620
QY 2143 ACACTGACTTCAAGAACCAACAACCTCTGTCTATCATTAACCTTGGCCAGGCTCTCAA 2202
DB 1621 ACACTGACTTCAAGAACCAACAACCTCTGTCTATCATTAACCTTGGCCAGGCTCTCAA 1680
QY 2203 GCTCAGCATTGGAGACCCAGCTCAGAGACTGTCTCTGTGCTGTGAGGAGCATCTGG 2262
DB 1681 GCTCAGCATTGGAGACCCAGCTCAGAGACTGTCTCTGTGCTGTGAGGAGCATCTGG 1740

QY 2263 CAAAAAAGACCTTTTCACTGCTCAGATGACCTTCAGAGAACATGAGTTAATGGGGCATC 2322
DB 1741 CAAAAAAGACCTTTTCACTGCTCAGATGACCTTCAGAGAACATGAGTTAATGGGGCATC 1800
QY 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGAGACACCCCATCTCTGAGCTCAGC 2382
DB 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGAGACACCCCATCTCTGAGCTCAGC 1860
QY 2383 TTCAATCACCTCTGTTTCAAGAGATTTTTCAGACAAATGTCTTGTGAGATGAG 2442
DB 1861 TTCAATCACCTCTGTTTCAAGAGATTTTTCAGACAAATGTCTTGTGAGATGAG 1920
QY 2443 AAGGAGAGAGTAAATCTTAATTCGATCATGATTTTGGAAAAAGCGCTAGAGAGATNT 2502
DB 1921 AAGGAGAGAGTAAATCTTAATTCGATCATGATTTTGGAAAAAGCGCTAGAGAGATNT 1980
QY 2503 GGAATACATGGCCTGTTTGGGGCATCAACACAGTTTCTATTTGGGCTGTAAATGAT 2562
DB 1981 GGAATACATGGCCTGTTTGGGGCATCAACACAGTTTCTATTTGGGCTGTAAATGAT 2040
QY 2563 GAGGGGAGAGAGATGAGAGAAATCTTTCACTGCTGGCTGTCTAGGGGAGAACTTG 2622
DB 2041 GAGGGGAGAGAGATGAGAGAAATCTTTCACTGCTGGCTGTCTAGGGGAGAACTTG 2100
QY 2623 ATGCAATGGGTCCCGTCCCTGAGCTGCTGCTGCAACCACTCTTGGAGTCCCTCAC 2682
DB 2101 ATGCAATGGGTCCCGTCCCTGAGCTGCTGCTGCAACCACTCTTGGAGTCCCTCAC 2160
QY 2683 TGCTTGTAGAGACTGGAGAACAAACGTTCTCTGACCAAGTGTGAGCCCATTCAGAGAA 2742
DB 2161 TGCTTGTAGAGACTGGAGAACAAACGTTCTCTGACCAAGTGTGAGCCCATTCAGAGAA 2220
QY 2743 ATGGGATGTGTGTAGAAAACAGCATGAGACTTTAGTGTGACCTTTCTGATTTAATTC 2802
DB 2221 ATGGGATGTGTGTAGAAAACAGCATGAGACTTTAGTGTGACCTTTCTGATTTAATTC 2280
QY 2803 AGCGGACAGTGAAGAAAGCTTCACTGATTTGAGGGAGAGACAGATTAATGAGAGC 2862
DB 2281 AGCGGACAGTGAAGAAAGCTTCACTGATTTGAGGGAGAGACAGATTAATGAGAGC 2340
QY 2863 CCCACATGATGATCTGCTGCTGAGTGGGTCCAGTCAAGATCTTAATGGCAGATTTCTC 2922
DB 2341 CCCACATGATGATCTGCTGCTGAGTGGGTCCAGTCAAGATCTTAATGGCAGATTTCTC 2400
QY 2923 TTCTCCGTCTCAAGGTCAACAGAAACCTGAAAGAGCTGACCTTAATGAGAACTGCTG 2982
DB 2401 TTCTCCGTCTCAAGGTCAACAGAAACCTGAAAGAGAGCTGACCTTAATGAGAACTGCTG 2460
QY 2983 AGCCACTGAGTGAAGAGTCTTGTAGAACCTGAGACGCTGAGCTGCTGAGAG 3042
DB 2461 AGCCACTGAGTGAAGAGTCTTGTAGAACCTGAGACGCTGAGCTGCTGAGAG 2520
QY 3043 ACCCTGCGTTGCTGCTGCTGCTGCTCAGAGCTGAGAGACTGAGAGACTTGTGGG 3102
DB 2521 ACCCTGCGTTGCTGCTGCTGCTGCTCAGAGCTGAGAGACTGAGAGACTTGTGGG 2580
QY 3103 CTGAGAGCCAACTGACACCTGACCGAGCTGAGCTTGAATGTCTCAAGATGCT 3162
DB 2581 CTGAGAGCCAACTGACACCTGACCGAGCTGAGCTTGAATGTCTCAAGATGCT 2640
QY 3163 GAGAGCAAAACACTTTGACAGAGACTGAGACGCGAGCTGCAAGGCTACAGGACCTGAG 3222
DB 2641 GAGAGCAAAACACTTTGACAGAGACTGAGACGCGAGCTGCAAGGCTACAGGACCTGAG 2700
QY 3223 CTGCTCAGCTGTGGCTCACTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
DB 2701 CTGCTCAGCTGTGGCTCACTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
QY 3283 AGCCCAAGCTTGAAGAGCTGAGCTGAGCAGAGAAACAACCTGATGATGCTTGGCGTGCAG 3342
DB 2761 AGCCCAAGCTTGAAGAGCTGAGCTGAGCAGAGAAACAACCTGATGATGCTTGGCGTGCAG 2820
QY 3343 CTGCTCTGTGAGGGGCTCAGGAGCTCTGCTGCAAACTATAGCGCTGAGGAGCTGAGACAG 3402

Db	2821	CTGCTCTGTAGAGGGGCTCAGGCACTCTGCTCGCAAACTCATACGCTGG-----	2869
QY	3403	ACAACTCTGAGTGAATGATGAGGCAAGAACTGAGGGCCCTGAGCAGAGAACTCAG	3462
Db	2870	-----	2869
QY	3463	CTGCTCATCTTCAGCACAAGGAAACCAAGTGTATGACCCCTTACTGAGGGCTTGATACG	3522
Db	2870	-----GGAAACCAAGTGTATGACCCCTTACTGAGGGCTTGATACG	2910
QY	3523	GGAGAGATGATATAGCACTCTCTCACTCAAGGGGCAAGACCTCGGATCAGAGGGCG	3582
Db	2911	GGAGAGATGATATAGCACTCTCTCACTCAAGGGGCAAGACCTCGGATCAGAGGGCG	2970
QY	3583	GCTTCCCATGTGTGCTCAGGCTTAATCTGAACTCTCGACGTGACCAAGATCTTCCCAAT	3642
Db	2971	GCTTCCCATGTGTGCTCAGGCTTAATCTGAACTCTCGACGTGACCAAGATCTTCCCAAT	3030
QY	3643	GCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGGAATCTTGTGCTGCT	3702
Db	3031	GCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGGAATCTTGTGCTGCT	3050
QY	3703	TTCCTCCGCTCTCAAGGGGACCTGCAATCAAGAGCTTTGGGACCTGACGATCTTCG	3762
Db	3091	TCTCTGCTCTCTCAAGGGGACCTGCAATCAAGAGCTTTGGGACCTGACGATCTTCG	3150
QY	3763	GGCCCCAAGGGGCTGTGCTACTAGAGGTAGTTACAAAGAAAGAACTTGACCGATT	3822
Db	3151	GGCCCCAAGGGGCTGTGCTACTAGAGGTAGTTACAAAGAAAGAACTTGACCGATT	3210
QY	3823	CACCTTCCCTGTAGCTGCTCCTCAACCGCTGACCAACGAGGTCTGTGCTTTGTATAGA	3882
Db	3211	CACCTTCCCTGTAGCTGCTCCTCAACCGCTGACCAACGAGGTCTGTGCTTTGTATAGA	3270
QY	3883	GAAGCGGTGACCGTGTGAGATTGAATTCGTGTGTGGGACCAAGTCCGTGGGAGATCAAC	3942
Db	3271	GAAGCGGTGACCGTGTGAGATTGAATTCGTGTGTGGGACCAAGTCCGTGGGAGATCAAC	3330
QY	3943	CCAACAGACAGCTGAGTGTGTGGCAAGGACCTCTGTCGACATCAAGGCTGAGCTCGAGCT	4002
Db	3331	CCAACAGACAGCTGAGTGTGTGGCAAGGACCTCTGTCGACATCAAGGCTGAGCTCGAGCT	3390
QY	4003	GTCGAGCTGTGACCTTCCCTCACTTTGTGCTCTCAAGGGGGCCATGTGACACATCC	4062
Db	3391	GTCGAGCTGTGACCTTCCCTCACTTTGTGCTCTCAAGGGGGCCATGTGACACATCC	3450
QY	4063	CTGTTCCAAATGCGCCACTTTAAAGAGAGGGGATGCTCTCTGAGAAAGCCAGCGGTG	4122
Db	3451	CTGTTCCAAATGCGCCACTTTAAAGAGAGGGGATGCTCTCTGAGAAAGCCAGCGGTG	3510
QY	4123	GAGCTGCATCATATGTTCTGGAAGAACCCCAAGCTTCTCCCTTGGGAGTCTCCTGAAA	4182
Db	3511	GAGCTGCATCATATGTTCTGGAAGAACCCCAAGCTTCTCCCTTGGGAGTCTCCTGAAA	3570
QY	4183	ATGATTCATAATGCGCCTGCGGCTTCAATTCGCGTCACTCTGTGTGTGCTTTACACGCG	4242
Db	3571	ATGATTCATAATGCGCCTGCGGCTTCAATTCGCGTCACTCTGTGTGTGCTTTACACGCG	3630
QY	4243	GTCATCTCTGAGGAAGTCACTTTCACCTCTTACTGATCCCAAGTGACTGCTCCATTGCG	4302
Db	3631	GTCATCTCTGAGGAAGTCACTTTCACCTCTTACTGATCCCAAGTGACTGCTCCATTGCG	3690
QY	4303	AAGGACCTGAGCTCTGCTATCGAAGCCCTGAGAAAGACCAAGCTGTTCTGGAGTTCTAC	4362
Db	3691	AAGGACCTGAGCTCTGCTATCGAAGCCCTGAGAAAGACCAAGCTGTTCTGGAGTTCTAC	3750
QY	4363	GTTGGCCACTTGGGATCAAGGATCAGGCTGCAAGTGAAGACCAAGAAAGATGAGACTTG	4422
Db	3751	GTTGGCCACTTGGGATCAAGGATCAGGCTGCAAGTGAAGACCAAGAAAGATGAGACTTG	3810
QY	4423	GTCGCGAGGCTTGTGAAACCAAGAGATCTCATGCTGTGAACTTACTCTGATCCCTCCA	4482

Db	3811	GTGTGGAGGGCTTTGGTGAATACAGAGAGATCTATGCTGTGCACACTACTGATCCCTCCA	3870
Qy	4483	GCCCGCATAGCCCGTACCTTCACTTGTGATGCCCCGACGTTGTGCACTTTGTGTGACAG	4542
Db	3871	GCCCGCATAGCCCGTACCTTCACTTGTGATGCCCCGACGTTGTGCACTTTGTGTGACAG	3930
Qy	4543	TATTCGAGAGCAGCTGATATGCCCGAGTGCATCCGTGAGAGGTTTGCTTGGACAAATCGCAT	4602
Db	3931	TATTCGAGAGCAGCTGATATGCCCGAGTGCATCCGTGAGAGGTTTGCTTGGACAAATCGCAT	3990
Qy	4603	GGAACAGGTGCTGAGCCACGAGACAGTACGAGAGGGTGTGCTGAGAAACAGAGGCCACAG	4662
Db	3991	GGAACAGGTGCTGAGCCACGAGACAGTACGAGAGGGTGTGCTGAGAAACAGAGGCCACAG	4050
Qy	4663	CAGATGCGAAGAGCTGTTCAAGCTTTGAGGCACTCTTGAGACCGGAAGTCCAAAGATGAGCTC	4722
Db	4051	CAGATGCGAAGAGCTGTTCAAGCTTTGAGGCACTCTTGAGACCGGAAGTCCAAAGATGAGCTC	4110
Qy	4723	TACCAAGCCTTGAAAGAGAGACCATTCCTCACTCATTTTGGAACTCTGGGAGAAAGGCGAGC	4782
Db	4111	TACCAAGCCTTGAAAGAGAGACCATTCCTCACTCATTTTGGAACTCTGGGAGAAAGGCGAGC	4170
Qy	4783	AAAAAGGAGCTCCTGCACCTCAGAGCTGA	4812
Db	4171	AAAAAGGAGCTCCTGCACCTCAGAGCTGA	4200

```

RESULT 10
US-09-388-221-5
; Sequence 5, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regulati
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4332)
; US-09-388-221-5

```

Query Match	72.7%	Score 3958	DB 10	Length 4332
Best Local Similarity	95.0%	Pred. No. 0		
Matches 4200	Conservative	0	Mismatches	0
			Indels	222
			Gaps	2
QY	523	ATGCTGGCGAGCGCTGGGGCCGCTGGCCCTGTTACTTGAAGTTCCTGAAGAAGAGAG	582	
Db	1	ATGGCTGGCGGAGCCTGGGGCCGCTGGCCCTGTTACTTGAAGTTCCTGAAGAAGAGAG	60	
QY	583	CTGAAGAGTTCCAGCTTCTGCTGCCAAATAAAGGCACTCCAGAGCTTTGGGTGAG	642	
Db	61	CTGAAGAGTTCCAGCTTCTGCTGCCAAATAAAGGCACTCCAGAGCTTTGGGTGAG	120	
QY	643	ACACCCGCTCAGCCAGAGAAACGAGTGGCATGAGAGTGCCCTGTACTTGGTGGCTCAG	702	
Db	121	ACACCCGCTCAGCCAGAGAAACGAGTGGCATGAGAGTGCCCTGTACTTGGTGGCTCAG	180	
QY	703	TATGGGGAGACGGGGCCTGGGACCTTAGCCCTCACTACTGGGAGCAGATGGGGCTGAGG	762	
Db	181	TATGGGGAGACGGGGCCTGGGACCTTAGCCCTCACTACTGGGAGCAGATGGGGCTGAGG	240	
QY	763	TCAGTGTGCGCCAAAGCCAGAGAAAGGGCAGGCCACTTCCTCTATTCCCTTACAGCCA	822	
Db	241	TCAGTGTGCGCCAAAGCCAGAGAAAGGGCAGGCCACTTCCTCTATTCCCTTACAGCCA	300	
QY	823	AGTGAACCCCACTGGGCTCTTCCAGCCCAACCACTTCAACCGCAGTGTAAATGCCCTGG	882	

Db 301 AGTGAACCCACCTGGGGTCTCCAGCAACCACTCCACCGAGTGAAGCCCTGG 360
Qy 883 ATTCATGAATTCGCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGAACAAGT 942
Db 361 ATTCATGAATTCGCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGAACAAGT 420
Qy 943 CCTGACATCTGGAAGCGCGCTGGAAGAAATTTCTGCTCACTCTCTCAACAAGCTCTT 1002
Db 421 CCTGACATCTGGAAGCGCGCTGGAAGAAATCTGCTCACTCTCTCAACAAGCTCTT 480
Qy 1003 CCAAGTCCCGACATGATGCTCCAGGCAAGGATCAACCAAGCCCAATCCACA 1062
Db 481 CCAAGTCCCGACATGATGCTCCAGGCAAGGATCAACCAAGCCCAATCCACA 540
Qy 1063 GCAGTCTGGGGAGCTGGGGATCCCACTCAGAGCCAGCTAGCAACCCAGAGAGAG 1122
Db 541 GCAGTCTGGGGAGCTGGGGATCCCACTCAGAGCCAGCTAGCAACCCAGAGAGAG 600
Qy 1123 GCTCTGGGACCCCAATGGCTCTGATGAAGCTCAGAAATTTACTACAGAAATCAGA 1182
Db 601 GCTCTGGGACCCCAATGGCTCTGATGAAGCTCAGAAATTTACTACAGAAATCAGA 660
Qy 1183 GAAAGAGAGAGAAATCAGAGAAAGCAGGCCCCATGGGCAAGCGGTGATGAAGC 1242
Db 661 GAAAGAGAGAGAAATCAGAGAAAGCAGGCCCCATGGGCAAGCGGTGATGAAGC 720
Qy 1243 CCCCACAGGCGCACACAGCCTTACAGCCCAACCAACCATAGGGAGCTTCTGAGA 1302
Db 721 CCCCACAGGCGCACACAGCCTTACAGCCCAACCAACCATAGGGAGCTTCTGAGA 780
Qy 1303 GAGAGCTCTGTTTCCATGAGCCCTGGAAAAATGAGAGATTTAACAAAAATTCACA 1362
Db 781 GAGAGCTCTGTTTCCATGAGCCCTGGAAAAATGAGAGATTTAACAAAAATTCACA 840
Qy 1363 CTGCTCTTCTCAAGAAGCTCAACCCAGAGCCAGATCCCTGTGCAAGAGAACTG 1422
Db 841 CTGCTCTTCTCAAGAAGCTCAACCCAGAGCCAGATCCCTGTGCAAGAGAACTG 900
Qy 1423 CCTGATTAATGAGAGAAATCAGAGCAATTTAATGAGATCAGAGATTTATTTGG 1482
Db 901 CCTGATTAATGAGAGAAATCAGAGCAATTTAATGAGATCAGAGATTTATTTGG 960
Qy 1483 GGGCTGATTCACCAAGAACTCGCATATGCTATCTGAGGGGGCTGTGGAATTTGG 1542
Db 961 GGGCTGATTCACCAAGAACTCGCATATGCTATCTGAGGGGGCTGTGGAATTTGG 1020
Qy 1543 TCAACACTGGGCAAGGAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGAC 1602
Db 1021 TCAACACTGGGCAAGGAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGAC 1080
Qy 1603 TTCCAGCATGCTTTCTACTTCAAGCTGAGAGAGCTGGCCCAATCAAGGTGATG 1662
Db 1081 TTCCAGCATGCTTTCTACTTCAAGCTGAGAGAGCTGGCCCAATCAAGGTGATG 1140
Qy 1663 GCTGAGCTATCGGAAAAAGTGGAGACCACTCCGGCTCCCAATTGACAGATCTGT 1722
Db 1141 GCTGAGCTATCGGAAAAAGTGGAGACCACTCCGGCTCCCAATTGACAGATCTGT 1200
Qy 1723 AGGCGAGAGGAGGCTCTTCACTCGATGATGTATGATGAGCCAGGATGGGCTTGG 1782
Db 1201 AGGCGAGAGGAGGCTCTTCACTCGATGATGTATGATGAGCCAGGATGGGCTTGG 1260
Qy 1783 GAGCCGAGTCTGAGCTCTGCTGCACTGAGCCAGCCAGCCAGCCAGGAGTGA 1842
Db 1261 GAGCCGAGTCTGAGCTCTGCTGCACTGAGCCAGCCAGCCAGCCAGGAGTGA 1320
Qy 1843 GGGAGTTTGGTGGGAAAAATTAATCTTCCGAGGACCTCTTCTGATCAAGGCTCG 1902
Db 1321 GGGAGTTTGGTGGGAAAAATTAATCTTCCGAGGACCTCTTCTGATCAAGGCTCG 1380
Qy 1903 ACAAGCTCTGAGAACTTCTCTTGTGAGAGGAGCAGTTGGGTGAGAGTCTGGGG 1962

Db 1381 ACAAGCTCTGAGAACTTCTCTTGTGAGAGGAGCAGCTTGGGTAGAGTCTGGGG 1440
Qy 1963 TTCTGATTCAGCAGAGAGAAATTTCTACAGATATTTTCAAGATGAAGAGCA 2022
Db 1441 TTCTGATTCAGCAGAGAGAAATTTCTACAGATATTTTCAAGATGAAGAGCA 1500
Qy 2023 ATTAGAGCTTTAGTTGGTCAATTAACAAAGAGCTCTGGGCTCTGTGTCTTGG 2082
Db 1501 ATTAGAGCTTTAGTTGGTCAATTAACAAAGAGCTCTGGGCTCTGTGTCTTGG 1560
Qy 2083 TGGGTGCTCTGGGCTCTGCACTTGGCTGATGACAGATGAAGGAGGAAAAATC 2142
Db 1561 TGGGTGCTCTGGGCTCTGCACTTGGCTGATGACAGATGAAGGAGGAAAAATC 1620
Qy 2143 ACACTGACTTCAAGACCAACAACCTCTGTCTTCAATTAACCTTGGCCAGGCTCCA 2202
Db 1621 ACACTGACTTCAAGACCAACAACCTCTGTCTTCAATTAACCTTGGCCAGGCTCCA 1680
Qy 2203 GCTCAGCCATTGGGACCCCAAGCTCAGAGACTCTGTCTCTGTGGCTCTGAGGCA 2262
Db 1681 GCTCAGCCATTGGGACCCCAAGCTCAGAGACTCTGTCTCTGTGGCTCTGAGGCA 1740
Qy 2263 CAAAAAAGACCTTTTCAAGTCAAGTGAAGCTCAGAGAGATGGGTTAGTGGGGCAT 2322
Db 1741 CAAAAAAGACCTTTTCAAGTCAAGTGAAGCTCAGAGAGATGGGTTAGTGGGGCAT 1800
Qy 2323 ATCTCCACCTTCTTGAAGATGGATATCTTCAAGAGACCCATCCCTGTGACTAC 2382
Db 1801 ATCTCCACCTTCTTGAAGATGGATATCTTCAAGAGACCCATCCCTGTGACTAC 1860
Qy 2383 TTCAATCACTCTGTTTCAAGAGTCTTGGACAAATGCTCTATCTTGGAGATGAG 2442
Db 1861 TTCAATCACTCTGTTTCAAGAGTCTTGGACAAATGCTCTATCTTGGAGATGAG 1920
Qy 2443 AAGGGAGAGTAAATCTTAATGATCATGATTTTGAAGAAAGCGCTAAGAGATAT 2502
Db 1921 AAGGGAGAGTAAATCTTAATGATCATGATTTTGAAGAAAGCGCTAAGAGATAT 1980
Qy 2503 GGAATACATGAGCTGTTGGGAGATCAACCAACGTTTCTATTTGGGCTGTTAATGAT 2562
Db 1981 GGAATACATGAGCTGTTGGGAGATCAACCAACGTTTCTATTTGGGCTGTTAATGAT 2040
Qy 2563 GAGGGGAGAGAGATGGAGAAACATTTTCACTGGGGCTGTCTCAGGGGAGAACTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAAACATTTTCACTGGGGCTGTCTCAGGGGAGAACTG 2100
Qy 2623 ATGAGTGGGTCCTGCTCCCTGAGCTGCTGAGCCACACTCTGTGAGTCCCTCAC 2682
Db 2101 ATGAGTGGGTCCTGCTCCCTGAGCTGCTGAGCCACACTCTGTGAGTCCCTCAC 2160
Qy 2683 TGCTTGACAGACTTGGAGAAACAAACGTTCTTGACAAAGTATGGCCCAATTCGA 2742
Db 2161 TGCTTGACAGACTTGGAGAAACAAACGTTCTTGACAAAGTATGGCCCAATTCGA 2220
Qy 2743 ATGGGATGATGTATGAACAAACATGAGCTCTTATGATGACATTTGCAATTAATC 2802
Db 2221 ATGGGATGATGTATGAACAAACATGAGCTCTTATGATGACATTTGCAATTAATC 2280
Qy 2803 AGCGGCAAGTGAAGAGCTTCAAGTGAAGGAGGAGGAGCAGATCAACATGAGAG 2862
Db 2281 AGCGGCAAGTGAAGAGCTTCAAGTGAAGGAGGAGGAGCAGATCAACATGAGAG 2340
Qy 2863 CCAACCATGATGATCTGTTCAAGTGGGATCCAGTCAAGATGCTTATTTGGAGATTCT 2922
Db 2341 CCAACCATGATGATCTGTTCAAGTGGGATCCAGTCAAGATGCTTATTTGGAGATTCT 2400
Qy 2923 TTCTCGCTCTCAAGGCTCAAGAACTGAAGAGCTGAGACTTAAGTGAAGAACTGCTG 2982
Db 2401 TTCTCGCTCTCAAGGCTCAAGAACTGAAGAGCTGAGACTTAAGTGAAGAACTGCTG 2460
Qy 2983 AGCACTCTGAGTGAAGAGCTTGTGAAGACCTTGAAGACGCTTGTGCTCTGAG 3042
Db 2461 AGCACTCTGAGTGAAGAGCTTGTGAAGACCTTGAAGACGCTTGTGCTCTGAG 2520

OY	3043	ACCCGCGGCTGGCGCTGGTGGCTGACAGCTGAGACGTGCAAGAGACCTTGCTTTGGG	3102
Db	2521	ACCTGCGGTTGGCTGGCTGTGGCTCAACAGCTGAGACGTCAAGAACCTTGGCTTTGGG	2580
OY	3103	CTGAGAGCCAAACCAAGACCTCTGACCGAGCTGAGACCTGAGCTTCAATTGTGCTCACAGGATGCT	3162
Db	2581	CTGAGAGCCAAACCAAGACCTCTGACCGAGCTGAGACCTTCAATTGTGCTCACAGGATGCT	2640
OY	3163	GGAGCCAAACACCTTTGGCAAGACTGAGACAGCGCGAGCTGCAAGCTACAGCGACTGCG	3222
Db	2641	GGAGCCAAACACCTTTGGCAAGACTGAGACAGCGCGAGCTGCAAGCTACAGCGACTGCG	2700
OY	3223	CTGGCTCAGCTGAGGCGCTCAGCGCTGACCTGCTGCGCCAGGACCTGGGCTGTGCTTAAGTGGC	3282
Db	2701	CTGGCTCAGCTGAGGCGCTCAGCGCTGACCTGCTGCGCCAGGACCTGGGCTGTGCTTAAGTGGC	2760
OY	3283	AGCCCCAGCCTGAAGAGAGCTAGACCTGCGAGAGAACCAACTGATGACGTTGGCGTGGCA	3342
Db	2761	AGCCCCAGCCTGAAGAGAGCTAGACCTGCGAGAGAACCAACTGATGACGTTGGCGTGGCA	2820
OY	3343	CTGCTCTGTGAGGGGCTCAGGCAATCTGCGCTGCAAACTCAATACGCTTGGGGCTGACCG	3402
Db	2821	CTGCTCTGTGAGGGGCTCAGGCAATCTGCGCTGCAAACTCAATACGCTTGG-----	2869
OY	3403	ACAACTCTGATGATGATGATGAGGACGAGGAACTGAGGGGCTTGGACAGAGAAACCTCAG	3462
Db	2870	-----	2869
OY	3463	CTGCTCATCTTCAAGCAGACGGAACCAAGTGTATGACCTCCTACTGAGGGCCTGATATCG	3522
Db	2870	-----GGAAACCAAGTGTATGACCTCCTACTGAGGGCCTGATATCG	2910
OY	3523	GGAGAGATGAGTAATATGACCAATCCCTCACTCAAGGGGCGAGACCTCGATCAAGAGGGCG	3582
Db	2911	GGAGAGATGAGTAATATGACCAATCCCTCACTCAAGGGGCGAGACCTCGATCAAGAGGGCG	2970
OY	3583	GCTTCCCAATGTTGCTCAGGCTAATCTCAACTCTCGAACGTGAGCAAGATCTTCCCAATT	3642
Db	2971	GCTTCCCAATGTTGCTCAGGCTAATCTCAACTCTCGAACGTGAGCAAGATCTTCCCAATT	3030
OY	3643	GCTGAGATGTCAGAGGAAAGCTCCCCCAAGGTAAGTACCGGTGGAACTTTGTGGTGGCT	3702
Db	3031	GCTGAGATGTCAGAGGAAAGCTCCCCCAAGGTAAGTACCGGTGGAACTTTGTGGTGGCT	3090
OY	3703	TCTCCCTGCTCTCAAGGGGACCTGCAATCGAAGCCTTTGGGGGACTGACGATGACTTCTGCG	3762
Db	3091	TCTCCCTGCTCTCAAGGGGACCTGCAATCGAAGCCTTTGGGGGACTGACGATGACTTCTGCG	3150
OY	3763	GGCCCCACGGGGGCTGTGGCTTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3822
Db	3151	GGCCCCACGGGGGCTGTGGCTTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3210
OY	3823	CACCTTCCCTGTAGCTGGCTCCATCCGCTGGGCCCAACAGGGTCACTGCTTTGTGTATGAGA	3882
Db	3211	CACCTTCCCTGTAGCTGGCTCCATCCGCTGGGCCCAACAGGGTCACTGCTTTGTGTATGAGA	3270
OY	3883	GAAAGCGTGAACGTTGAGATTGAATCTGTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC	3942
Db	3271	GAAAGCGTGAACGTTGAGATTGAATCTGTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC	3330
OY	3943	CCACAGCACAACCTGAGTGTGTGGGAGGGGCTCTGCTGACATCAAGAGCTGACCTGAGACT	4002
Db	3331	CCACAGCACAACCTGAGTGTGTGGGAGGGGCTCTGCTGACATCAAGAGCTGACCTGAGACT	3390
OY	4003	GTCGAGCTGTGACCTTCCCTCACTTTGTGGCTCTTCAAGGGGGCCATGTGAGACATCC	4062
Db	3391	GTCGAGCTGTGACCTTCCCTCACTTTGTGGCTCTTCAAGGGGGCCATGTGAGACATCC	3450
OY	4063	CTGTTCCTCAATGGCCCACTTTAAAGAGAGGGGATGCTCTGTGAGAGACCGACCGGGTG	4122
Db	3451	CTGTTCCTCAATGGCCCACTTTAAAGAGAGGGGATGCTCTGTGAGAGACCGACCGGGTG	3510

QY	4123	GGCGTCATCATATGTTCTGGAAAAACCCAGCTTCTCCCCCTGGAGAGCTCCGAAA	4182
Db	3511	GAGCTGCATCATATGTTCTGGAAAAACCCAGCTTCTCCCCCTGGAGAGCTCCGAAA	3570
QY	4183	ATGATCCATAATGACCTCGCGCTTCATTCCTCCGACCTCTGTGGTGTGTGCTTTACACCGC	4242
Db	3571	ATGATCCATAATGACCTCGCGCTTCATTCCTCCGACCTCTGTGGTGTGTGCTTTACACCGC	3630
QY	4243	GTCCATCTGAGGAAGTCACCTTCCACTCTACCTGATCCCAAGTACCTGCTCCATTCCG	4302
Db	3631	GTCCATCTGAGGAAGTCACCTTCCACTCTACCTGATCCCAAGTACCTGCTCCATTCCG	3690
QY	4303	-----	4302
Db	3691	AAGCCATAGATGATCTAGAAAAGAAATTCAGTTTGTGCGAATCCACAAAGCACC	3750
QY	4303	-----	4302
Db	3751	CTGACCCCACTTATATGAGGCTGTGTGTTACCTGTGTCTGGGCTGTGGTTCAAGGATGCTG	3810
QY	4303	-----AAGGAATGAGGCTCTGTCTATTCGAAGCCCTGGAGAGACCAAGCTGTTC	4350
Db	3811	GAAATACTCCCCAAGAACTGAGGCTGTGTCTATTCGAAGCCCTGGAGAGACCAAGCTGTTC	3870
QY	4351	TCGGAGTTCTACGTTGGGACCACTTGGGATCAGGGAATCAGGCTGCAAGTAAAGA	4410
Db	3871	TCGGAGTTCTACGTTGGGACCACTTGGGATCAGGGAATCAGGCTGCAAGTAAAGA	3930
QY	4411	GATGAGACTCTGTGTGTGGAGGCTTGTGTAAACAGAGATCTCATCTCTGCACTACT	4470
Db	3931	GATGAGACTCTGTGTGTGGAGGCTTGTGTAAACAGAGATCTCATCTCTGCACTACT	3990
QY	4471	CTGATCTCCTCCAGCCCCGATAGCCGTAACCTTCACTCTGTGATGCCCCGCACTGTGACAC	4530
Db	3991	CTGATCTCCTCCAGCCCCGATAGCCGTAACCTTCACTCTGTGATGCCCCGCACTGTGACAC	4050
QY	4531	TTTGTGAGCAAGTATTCGAGAGAGAGCTGATPAGCCCGAGTGAACATCGGATGAGGTTGCTTG	4590
Db	4051	TTTGTGAGCAAGTATTCGAGAGAGAGCTGATPAGCCCGAGTGAACATCGGATGAGGTTGCTTG	4110
QY	4591	GACAAACTGCATGACAGGTGTGTGAGCCAGAGACAGTACGAGAGGGTGTGCTGTGAGAAC	4650
Db	4111	GACAAACTGCATGACAGGTGTGTGAGCCAGAGACAGTACGAGAGGGTGTGCTGTGAGAAC	4170
QY	4651	ACGAGGCCACGCAATGCGGAAGCTGTTCACTTGAAGCGAAGCTGTGGGACCGGAAGTGC	4710
Db	4171	ACGAGGCCACGCAATGCGGAAGCTGTTCACTTGAAGCGAAGCTGTGGGACCGGAAGTGC	4230
QY	4711	AAAGATGAGCTCTACCAAGCCCTGAAGAGAGACCCATCTCTCACTCTTATGAACTCTGG	4770
Db	4231	AAAGATGAGCTCTACCAAGCCCTGAAGAGAGACCCATCTCTCACTCTTATGAACTCTGG	4290
QY	4771	GAGAAAGGCAAGAAAGGAACTCTCTGCTCACTCAGCACTGA	4812
Db	4291	GAGAAAGGCAAGAAAGGAACTCTCTGCTCACTCAGCACTGA	4332

```

RESULT 11
; US-09-388-221-9
; Sequence 9, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regulat
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence

```



```

?   FEATURE:
?   NAME/KEY: CDS
?   LOCATION: (1)..(4362)
?   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
?   OTHER INFORMATION: Construct
US-09-388-221-9

```

Query Match	62.5%	Score 3400.4	DB 10	Length 4556
Best Local Similarity	93.7%	Pred. No. 0		
Matches 355	Conservative	0	Mismatches 236	Indels 3
				Gaps 1

QY	523	ATGGCTGGCGGAGGCTGGGGGCGGGCTGGGCTGGTACTTGGAGATTCCTGAAGAAGAGGAG	582
Db	1	ATGGCTGGCGGAGGCTGGGGGCGGGCTGGGCTGGTACTTGGAGATTCCTGAAGAAGAGGAG	60
QY	583	CTGAAGAAGTTCACAGCTTCTGCTGGCCATTAAGGCGCATCCACAGAGCTTTCGGGTGAG	642
Db	61	CTGAAGAAGTTCACAGCTTCTGCTGGCCATTAAGGCGCATCCACAGAGCTTTCGGGTGAG	120
QY	643	ACACCCGCTCAGCCAGAGAAGAAGAGTGGCATGAGAGGGGCGTGTACCTGGTGGCTCAG	702
Db	121	ACACCCGCTCAGCCAGAGAAGAAGAGTGGCATGAGAGGGGCGTGTACCTGGTGGCTCAG	180
QY	703	TATGGGAGAGCAGGGGCTGGAGCCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG	762
Db	181	TATGGGAGAGCAGGGGCTGGAGCCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG	240
QY	763	TCACTGTGGCCCAAGCCAGAGAAGGGCAGGGCATCTCCCTCATTCCTTACAGCCCA	822
Db	241	TCACTGTGGCCCAAGCCAGAGAAGGGCAGGGCATCTCCCTCATTCCTTACAGCCCA	300
QY	823	AGTGAACCCCAACGAGGGGTCTCCAGGCGCAACCCACCTCCACCGGAGTCAATAGCCCTGG	882
Db	301	AGTGAACCCCAACGAGGGGTCTCCAGGCGCAACCCACCTCCACCGGAGTCAATAGCCCTGG	360
QY	883	ATCCATGAATTCGCGGCGGGGTGACCCAGGGGCTCAGAGAGAAGGTTTTGAGACAGCTG	942
Db	361	ATCCATGAATTCGCGGCGGGGTGACCCAGGGGCTCAGAGAGAAGGTTTTGAGACAGCTG	420
QY	943	CTTGACACATCTGGAGCGCCGCTGGAGAGAATCTCTGCTCACTCTTACCAAGCTCTT	1002
Db	421	CTTGACACATCTGGAGCGCCGCTGGAGAGAATCTCTGCTCACTCTTACCAAGCTCTT	480
QY	1003	CCAAGCTCCCCAGACCAATGAGTCTCCAAAGCAGGAGTACCCACAGGCCCCACATCCACA	1062
Db	481	CCAAGCTCCCCAGACCAATGAGTCTCCAAAGCAGGAGTACCCACAGGCCCCACATCCACA	540
QY	1063	GCAGTGCTGGGGAGCTGGGAGTCCCCACCTCAGGCCAGCCTTAGCACCCAAGAGCAGAGG	1122
Db	541	GCAGTGCTGGGGAGCTGGGAGTCCCCACCTCAGGCCAGCCTTAGCACCCAAGAGCAGAGG	600
QY	1123	GCTCTCTGGGACCCCAATGGCCCTCTGGATGAACGTCAGGAATTTCTACAGAAATTCAGA	1182
Db	601	GCTCTCTGGGACCCCAATGGCCCTCTGGATGAACGTCAGGAATTTCTACAGAAATTCAGA	660
QY	1183	GAAAGAGAGAGAGAATCAGAGAAAGGCGAGGCCCATGGGAGCGGCTGTAGAAAGC	1242
Db	661	GAAAGAGAGAGAGAATCAGAGAAAGGCGAGGCCCATGGGAGCGGCTGTAGAAAGC	720
QY	1243	CCCCCAAGGGCGCACACCAGCTTACAGTCCCAACAACCCATGGAGGCTTCTGTGAGA	1302
Db	721	CCCCCAAGGGCGCACACCAGCTTACAGTCCCAACAACCCATGGAGGCTTCTGTGAGA	780
QY	1303	GAGAGCGCTGGTTCACATGGCCCTGGAGAAAATAGAGTTTAAACAAAATTCACACAG	1362
Db	781	GAGAGCGCTGGTTCACATGGCCCTGGAGAAAATAGAGTTTAAACAAAATTCACACAG	840
QY	1363	CTGCTACTTTCACAAAGACTCACCCCAAGACCAAGATTCCTGCTTCAGAGAGAGCTGG	1422
Db	841	CTGCTACTTTCACAAAGACTCACCCCAAGACCAAGATTCCTGCTTCAGAGAGAGCTGG	900
QY	1423	CCGATATTATGGAGAGAAATCAGAGACATTTAATTGAGATCAGAGACTTATTGGCCCA	1482

Dp	901	CTGATTAATGAGGAGGAATTCGAGCACTTAAATGAGATCAGAGCACTTAATTGGCCCA	960
Qy	1483	GGCTCGATATCCGAGAACTCCGATATGTCACTATCTCAGGGGGCTGCTGGAAATGGAG	1542
Dp	961	GGCTCGATATCCGAGAACTCCGATATGTCACTATCTCAGGGGGCTGCTGGAAATGGAG	1020
Qy	1543	TCAACATGGCCAGGGCAGGTGGAAGAAAGCCGTGGGGGAGAGGGCCAGCTGTATGGGGAACCG	1602
Dp	1021	TCAACATGGCCAGGGCAGGTGGAAGAAAGCCGTGGGGGAGAGGGCCAGCTGTATGGGGAACCG	1080
Qy	1603	TTCCAGCATGTCTTCTACTTCAAGCTGAGAGAGCTGGCCCAAGTCCAAGGTGTAGTCTC	1662
Dp	1081	TTCCAGCATGTCTTCTACTTCAAGCTGAGAGAGCTGGCCCAAGTCCAAGGTGTAGTCTC	1140
Qy	1663	GCTGAGCTCATTCGAAAAAATGGGACAGCCATCTCCGCTCCCATTTAGACAGATCTCTCT	1722
Dp	1141	GCTGAGCTCATTCGAAAAAATGGGACAGCCATCTCCGCTCCCATTTAGACAGATCTCTCT	1200
Qy	1723	AGGCGAAGAGGGCGTCTCTCATCTCTGAGATGTGTATATAGCCAGATGGCTTTCCAG	1782
Dp	1201	AGGCGAAGAGGGCGTCTCTCATCTCTGAGATGTGTATATAGCCAGATGGCTTTCCAG	1260
Qy	1783	GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGGGGGATGCACTGCTG	1842
Dp	1261	GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGGGGGATGCACTGCTG	1320
Qy	1843	GGCAGTTTGTGGGAAAAATATATCTTCCGAGGCATCTTCTGATCAGCGCTGGACC	1902
Dp	1321	GGCAGTTTGTGGGAAAAATATATCTTCCGAGGCATCTTCTGATCAGCGCTGGACC	1380
Qy	1903	ACAGCTCTGCAAAACCTCATTTCTTCTTTGGAGCAGGCACGTTGGGTAGAGTCTTGGG	1962
Dp	1381	ACAGCTCTGCAAAACCTCATTTCTTCTTTGGAGCAGGCACGTTGGGTAGAGTCTTGGG	1440
Qy	1963	TTCTCTGAGTCCAGCAGAAAGAAATATTTCTACAGATTTTCAACAGATGAAAGGCAAGCA	2022
Dp	1441	TTCTCTGAGTCCAGCAGAAAGAAATATTTCTACAGATTTTCAACAGATGAAAGGCAAGCA	1500
Qy	2023	ATTAGAGCTTTAGTGTGTCAATTCAAACAAAGAGCTTGCGCCCTGTGTCTTGTGCC	2082
Dp	1501	ATTAGAGCTTTAGTGTGTCAATTCAAACAAAGAGCTTGCGCCCTGTGTCTTGTGCC	1560
Qy	2083	TGGGTGTCTGGCTGGGCTGTGCACTTGGCTGATGTGACAGATGAAAGCCGAAGGAAAACTC	2142
Dp	1561	TGGGTGTCTGGCTGGGCTGTGCACTTGGCTGATGTGACAGATGAAAGCCGAAGGAAAACTC	1620
Qy	2143	ACACTGACTTCCAAAGACACACACAACACCTCGTCTACATTAACCTTGGCCAGGCTCTCCAA	2202
Dp	1621	ACACTGACTTCCAAAGACACACACAACACCTCGTCTACATTAACCTTGGCCAGGCTCTCCAA	1680
Qy	2203	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTTGCTCTGTGGCTGTGAGGGCATGTGG	2262
Dp	1681	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTTGCTCTGTGGCTGTGAGGGCATGTGG	1740
Qy	2263	CAAAAAAAGACCTTTTCAATCCAGATGACCTCAGAGAGCATGGGTTAAGTGGGGCCATC	2322
Dp	1741	CAAAAAAAGACCTTTTCAATCCAGATGACCTCAGAGAGCATGGGTTAAGTGGGGCCATC	1800
Qy	2323	ATCTCCACCTTCTTGAAGATGGGGAATCTTCAAGAGACCCCAATCCCTCGAGCTACAGC	2382
Dp	1801	ATCTCCACCTTCTTGAAGATGGGGAATCTTCAAGAGACCCCAATCCCTCGAGCTACAGC	1860
Qy	2383	TTCAATTCACCTGTGTTCGAAAGGTTCTTGGACGACATGTCTATGTCTTGGAGATGAG	2442
Dp	1861	TTCAATTCACCTGTGTTCGAAAGGTTCTTGGACGACATGTCTATGTCTTGGAGATGAG	1920
Qy	2443	AAGGGGAGAGGTAAACATTCTAATTGCATATAGATTGGAAAAAGACGCTAAGAACATAT	2502
Dp	1921	AAGGGGAGAGGTAAACATTCTAATTGCATATAGATTGGAAAAAGACGCTAAGAACATAT	1980
Qy	2503	GGAAATACATGGCCGTTTGGGGGCAATCAACACAGTTTCCATTTGGGCTGTTAAGAT	2562
Dp	1981	GGAAATACATGGCCGTTTGGGGGCAATCAACACAGTTTCCATTTGGGCTGTTAAGAT	2040

Db 9 GAAGAAGCTTCACTGATTGAGGCGAGGACAGATCAACATGAGAGCCCAACCATGCT 68
 QY 2874 AGTCTGTTCAAGTGGGTCCAGTCAAGATGCTATTTGGAGAGATTCTTCTCCGCTCT 2933
 Db 69 AGTCCGTTCAGGTGGGTCCAGTCAAGATGCTATTTGGAGAGATTCTTCTCCGCTCT 128
 QY 2934 CAAGGTCAACCAAAAGCTGAGAGAGCTTAAAGTGAAGAACTCGCTGAGCACTCTGC 2993
 Db 129 CAAGGTCAACCAAAAGCTGAGAGAGCTTAAAGTGAAGAACTCGCTGAGCACTCTGC 188
 QY 2994 AGTGAAGAGCTTTTGAAGACCTGAGAGAGCTTCTGCTGCTCTGAGAGACCTGGGTT 3053
 Db 189 AGTGAAGAGCTTTTGAAGACCTGAGAGAGCTTCTGCTGCTCTGAGAGACCTGGGTT 248
 QY 3054 GGTGCTGTGCTTCAAGCTGAGAGAGCTTCAAGAGACCTTTCCTTTGGCTGAGAGCA 3113
 Db 249 GGTGCTGTGCTTCAAGCTGAGAGAGCTTTCCTTTGGCTGAGAGCA 308
 QY 3114 CCAGACCTTGAACCTGAGAGAGCTTCAAGTGTCTTCAAGAGAGCTTGAAGAGCA 3173
 Db 309 CCAGACCTTGAACCTGAGAGAGCTTCAAGTGTCTTCAAGAGAGCTTGAAGAGCA 368
 QY 3174 CTTTTCAGAGAGCTGAGAGAGCTGAGAGAGCTTCAAGAGAGCTGAGAGAGCTGAG 3233
 Db 369 CTTTTCAGAGAGCTGAGAGAGCTGAGAGAGCTTCAAGAGAGCTGAGAGAGCTGAG 428
 QY 3234 TGGCTTCACTGTGAGCTGAGAGAGCTTGGCTTGTGTGAGAGAGCTTCAAGAGAGCT 3293
 Db 429 TGGCTTCACTGTGAGCTGAGAGAGCTTGGCTTGTGTGAGAGAGCTTCAAGAGAGCT 488
 QY 3294 GAAGAGCTTGAACCTGAGAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAG 3353
 Db 489 GAAGAGCTTGAACCTGAGAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAG 548
 QY 3354 GGGGCTCAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAG 3413
 Db 549 GGGGCTCAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAGAG 607
 QY 3414 TGATGATGAGAGAGAGAGCTTGAAGAGAGCTTGAAGAGAGAGCTTGAAGAGAGAG 3473
 Db 608 TGATGATGAGAGAGAGAGCTTGAAGAGAGCTTGAAGAGAGAGCTTGAAGAGAGAG 667
 QY 3474 CAGCAGAGAGAGAGAGAGCTTGAAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAG 3533
 Db 668 CAGCAGAGAGAGAGAGAGCTTGAAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAG 726
 QY 3534 TAATAGACAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAG 3593
 Db 727 TAATAGACAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAG 786
 QY 3594 TGTCTAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAG 3653
 Db 787 TGTCTAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAG 846
 QY 3654 AGAGAGAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAG 3713
 Db 847 AGAGAGAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAG 906
 QY 3714 TCAAGAGAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAG 3773
 Db 907 TCAAGAGAGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAG 966
 QY 3774 GCTGTGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAG 3833
 Db 967 GCTGTGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAGAG 1026
 QY 3834 AGCTGTGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAGAG 3893
 Db 1027 AGCTGTGAGCTTCAAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAGAG 1086
 QY 3894 CGTTGAGATGAGATTTCTGT 3953
 Db 1087 CGTTGAGATGAGATTTCTGT 1146

QY 3954 CTGATGTGTGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4013
 Db 1147 CTGATGTGTGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1206
 QY 4014 GACCTTCCCTCACTTTTGT 4073
 Db 1207 GACCTTCCCTCACTTTTGT 1266
 QY 4074 GAGCCCACTTTAAAGAGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4133
 Db 1267 GAGCCCACTTTAAAGAGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1326
 QY 4134 CATAGTTCTGAGAGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4193
 Db 1327 CATAGTTCTGAGAGAGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1386
 QY 4194 TGGCTTGGCTTCACTTCCGATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4253
 Db 1387 TGGCTTGGCTTCACTTCCGATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1446
 QY 4254 GGAAGTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 4313
 Db 1447 GGAAGTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1506
 QY 4314 GCTTGTCTATGAG 4373
 Db 1507 GCTTGTCTATGAG 1566
 QY 4374 GGAATCAG 4433
 Db 1567 GGAATCAG 1626
 QY 4434 CTTGTGTAAACAG 4493
 Db 1627 CTTGTGTAAACAG 1686
 QY 4494 GGTACCTTCACTTGT 4553
 Db 1687 GGTACCTTCACTTGT 1746
 QY 4554 GCTGTATGAG 4613
 Db 1747 GCTGTATGAG 1806
 QY 4614 GAGCCAG 4673
 Db 1807 GAGCCAG 1866
 QY 4674 GCTTGTCACTTGT 4733
 Db 1867 GCTTGTCACTTGT 1926
 QY 4734 GAAAG 4793
 Db 1927 GAAAG 1985
 QY 4794 CTTGCCACTCAG 4853
 Db 1986 CTTGCCACTCAG 2045
 QY 4854 CTGACCTTCTTGT 4913
 Db 2046 CTGACCTTCTTGT 2105
 QY 4914 TCCAG 4973
 Db 2106 TCCAG 2165
 QY 4974 AGGAGATCCAG 5033
 Db 2166 AGGAGATCCAG 2225

QY 5034 TGAGCTGCGAAGACCCCTGAGACCTCATAGAGCTCATCTGGGCGCACAGGCCAA 5093
 DB 2226 TGGAGCTGGCAAGACCCCTGAGACCTCATAGAGCTCATCTGGGCGCACAGGCCAA 2285
 QY 5094 GCCTAGAGCCCTCCGAGATCCCATCCAGCGCAAGAGAAATGAGAGGAGACATGAGACCAT 5153
 DB 2286 AGCTTAGAGCCCTCCGAGATCCCATCCAGCGCAAGAGAAATGAGAGGAGACATGAGACCAT 2345
 QY 5154 TTGGCTCTGGCTGTGTCTACAGGGGTGAGGCCCAAAATTTGGGGTTCAAGCTGGGAGGCCACG 5213
 DB 2346 TTGGCTCTGGCTGTGTCTACAGGGGTGAGGCCCAAAATTTGGGGTTCAAGCTGGGAGGCCACG 2405
 QY 5214 TGAATCTTGGCTTTGTACAGAGAAATCTACAGAGCAAGCCAAACAGAGTAAAGTGAAG 5273
 DB 2406 TGAATCTTGGCTTTGTACAGAGAAATCTACAGAGCAAGCCAAACAGAGTAAAGTGAAG 2465
 QY 5274 GAAGTTATTCAGAAATTAAGAGATACAGAGCTCTTTAGAAATTTGTCTAGCAGGCTT 5333
 DB 2466 GAAGTTATTCAGAAATTAAGAGATACAGAGCTCTTTAGAAATTTGTCTAGCAGGCTT 2525
 QY 5334 TCCAGTTTTCACGAAAGACCCCTTAATTAATTTTAAATTTTAAATTTAAAGATTA 5393
 DB 2526 TCCAGTTTTCACGAAAGACCCCTTAATTAATTTTAAATTTTAAATTTAAAGATTA 2585
 QY 5394 AAAAATACAAAAAAGAAAAATGAAGAAATTAAGAAATTAAGAGTTAAGTAACTAC 5444
 DB 2586 AAAAATACAAAAAAGAAAAATGAAGAAATTAAGAAATTAAGAGTTAAGTAACTAC 2636

RESULT 14

US-09-956-712-10
 ; Sequence 10, Application US/09956712
 ; Publication No. US20030092648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan M. Freiler
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
 ; FILE REFERENCE: RTS-0326
 ; CURRENT APPLICATION NUMBER: US/09/956,712
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 10
 ; LENGTH: 96649
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(96649)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-956-712-10

Query Match 31.4%; Score 1710.8; DB 11; Length 96649;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1712; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1168 TACACAGAAATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGACGAGCCCACTGGGCA 1227
 DB 30558 TACTTAGAAATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGACGAGCCCACTGGGCA 30617
 QY 1228 GCGGTGAGAGAAAGCCCAAGGCGCACACAGCTTACAGAGCCCAACACCAACCATGG 1287
 DB 30618 GCGGTGAGAGAAAGCCCAAGGCGCACACAGCTTACAGAGCCCAACACCAACCATGG 30677
 QY 1288 GAGCTTCTGTGAGAGAGAGGCTCTGTTCACATAGCCCTGAGAAATGAGAGATTTTAA 1347
 DB 30678 GAGCTTCTGTGAGAGAGAGGCTCTGTTCACATAGCCCTGAGAAATGAGAGATTTTAA 30737
 QY 1348 CAAATATTCACAGAGCTCTCTCTACAAAGACCTCACCCCAAGGCAAGATCCCTG 1407
 DB 30738 CAAATATTCACAGAGCTCTCTCTACAAAGACCTCACCCCAAGGCAAGATCCCTG 30797
 QY 1408 GTCAAGAGAGCTGCTGATTAATGTGAGAGAGATCAGAGCATTTAATGAGATCAGA 1467

DB 30798 GTCAAGAGAGCTGCTGATTAATGTGAGAGAGATCAGAGCATTTAATGAGATCAGA 30857
 QY 1468 GACTTATTTGGCCCAAGGCTTGATACCAAGAACTTGCAATGTCATCTGACAGGGGCT 1527
 DB 30858 GACTTATTTGGCCCAAGGCTTGATACCAAGAACTTGCAATGTCATCTGACAGGGGCT 30917
 QY 1528 GCTGGAATTTGGGAAGTCAACATCTGGCCAGGCTGAGGAAGCTGGGGGAGAGGCGAG 1587
 DB 30918 GCTGGAATTTGGGAAGTCAACATCTGGCCAGGCTGAGGAAGCTGGGGGAGAGGCGAG 30977
 QY 1588 CTGTATGGGAGACGCTTCCAGACATGTCTTCTACTTCACTGACAGAGAGCTGGCCAGTCC 1647
 DB 30978 CTGTATGGGAGACGCTTCCAGACATGTCTTCTACTTCACTGACAGAGAGCTGGCCAGTCC 31037
 QY 1648 AAGGTGTGAGTCTCGCTGAGCTCATGTGAAGAGTGGACAGCCACTCCGGCTCCATT 1707
 DB 31038 AAGGTGTGAGTCTCGCTGAGCTCATGTGAAGAGTGGACAGCCACTCCGGCTCCATT 31097
 QY 1708 AGACAGATCCGTGTAGAGCCAGAGCGGCTGCTTCACTCCGATGGTGTGATGAGCCA 1767
 DB 31098 AGACAGATCCGTGTAGAGCCAGAGCGGCTGCTTCACTCCGATGGTGTGATGAGCCA 31157
 QY 1768 GATGGGCTTGGCAGAGCCGAGTTCTGAGCTGTCTGCACTGAGCCAGCCAGCGCG 1827
 DB 31158 GATGGGCTTGGCAGAGCCGAGTTCTGAGCTGTCTGCACTGAGCCAGCCAGCGCG 31217
 QY 1828 GCGGATCACTGCTGGGCAAGTTCTGGGAGAAATCTACTCTCCGAGGATCCTTCTG 1887
 DB 31218 GCGGATCACTGCTGGGCAAGTTCTGGGAGAAATCTACTCTCCGAGGATCCTTCTG 31277
 QY 1888 ATCAGGCTGGGACCAAGCTCTGCAAGAACTCACTTCTTTGGAGCGGACGTTGG 1947
 DB 31278 ATCAGGCTGGGACCAAGCTCTGCAAGAACTCACTTCTTTGGAGCGGACGTTGG 31337
 QY 1948 GTAGAGCTCTGGGGTCTCTGAGTCCAGAGAGAGAAATTTCTACAGATATTTTACA 2007
 DB 31338 GTAGAGCTCTGGGGTCTCTGAGTCCAGAGAGAGAAATTTCTACAGATATTTTACA 31397
 QY 2008 GATGAAGGCAAGCAATTAAGACCTTTAGTTGTTGTCATTAATCAAAAGAGCTTGGGCC 2067
 DB 31398 GATGAAGGCAAGCAATTAAGACCTTTAGTTGTTGTCATTAATCAAAAGAGCTTGGGCC 31457
 QY 2068 CTGTGCTTGTGCGCTGGGAGTCTGCTGGCTGGCACTTGCCGATGAGCAGATGAAG 2127
 DB 31458 CTGTGCTTGTGCGCTGGGAGTCTGCTGGCTGGCACTTGCCGATGAGCAGATGAAG 31517
 QY 2128 CGGAAGGAAAACTCACATGACTTCCAAAGACACACAAACCTCTGTCTTACATTACCTT 2187
 DB 31518 CGGAAGGAAAACTCACATGACTTCCAAAGACACACAAACCTCTGTCTTACATTACCTT 31577
 QY 2188 GCCAGGCTCTCCAGGCTCAAGCAATTTGGACCCCAAGCTCAAGAGCTCTGCTCTGACT 2247
 DB 31578 GCCAGGCTCTCCAGGCTCAAGCAATTTGGACCCCAAGCTCAAGAGCTCTGCTCTGACT 31637
 QY 2248 GCTGAGGGCAATCTGGCAAAAAGACCTTTTCAAGTCAAGATGACCTCAGGAACATGGG 2307
 DB 31638 GCTGAGGGCAATCTGGCAAAAAGACCTTTTCAAGTCAAGATGACCTCAGGAACATGGG 31697
 QY 2308 TTAGATGGGCAATCATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGCAACCCATC 2367
 DB 31698 TTAGATGGGCAATCATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGCAACCCATC 31757
 QY 2368 CTTCTGAGCTACAGCTTCATTCACCTCTGTCTTCCAAAGATTTCTTGAAGCAATGCTCAT 2427
 DB 31758 CTTCTGAGCTACAGCTTCATTCACCTCTGTCTTCCAAAGATTTCTTGAAGCAATGCTCAT 31817
 QY 2428 GTCTTGAAGATGAGAGGGGAGAGTAAACATTTAATGCAATCATAAGATTGGAAG 2487
 DB 31818 GTCTTGAAGATGAGAGGGGAGAGTAAACATTTAATGCAATCATAAGATTGGAAG 31877
 QY 2488 ACGCTAGAGCATATGATATACATGAGCTGTTTGGGCAATCAACACAGTTTCTATTG 2547
 DB 31878 ACGCTAGAGCATATGATATACATGAGCTGTTTGGGCAATCAACACAGTTTCTATTG 31937

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:06 ; Search time 1342.51 Seconds
(without alignments)
10946.442 Million cell updates/sec

Title: US-09-996-617-1

Perfect score: 5444
Sequence: 1 gcccccggcgctggagagtc.....aggataagaagttaccac 5444

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA2001A.DAT.*
24: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneq-n-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5444	100.0	5523	22	AAF83651 Human CARD-7 polyp
2	5444	100.0	5523	25	ABSS5497 CDNA encoding huma
3	5439.2	99.9	5523	25	ABSS5497 Pyrin domain conta
4	4879.2	89.6	5100	24	AAI47127 Human G-protein co
5	4765.6	87.5	5122	22	AAI12951 Human NB-ARC and C
6	4148	76.2	4422	22	AAI02760 Human NB-ARC and C
7	4078	74.9	4194	22	AAI02761 Human NB-ARC and C
8	3952	72.6	4329	22	AAI02762 Human NB-ARC and C

9	3400.4	62.5	4556	22	AAI02764 Human NAC beta iso
10	3210.4	59.0	4466	22	AAI02765 Human NAC gamma or
11	2572.6	47.3	2657	21	AAI78392 Human secreted pro
12	484.4	8.9	578	22	AAI01744 Human reproductive
13	484.4	8.9	578	22	ABV97037 Human testicular a
14	482.8	8.9	487	24	ABV97880 Human pancreatic c
15	414.4	7.6	416	24	ABV97891 Human pancreatic c
16	399.8	7.3	3186	24	AAI44363 Human pyrin-8 CDNA
17	398	7.3	3300	24	AAI47129 Human pyrin domain conta
18	369.8	6.8	3857	22	AAI43233 Human PYRIN 1 (PYR
19	369.8	6.8	3857	25	ABX93556 Huma CDNA encoding
20	357.2	6.6	2763	24	ABO75801 Human MDT-13 enco
21	319.6	5.9	363	21	AAI08520 Human secreted pro
22	316.2	5.8	330	21	AACT5514 Human ORFX ORF1069
23	313.4	5.8	2691	24	AAI44365 Human PYRIN-11 CDN
24	304.6	5.6	2847	22	AAI01487 Human secreted pro
25	304.6	5.6	2847	25	ABZ73494 Secreted protein-e
26	255.2	4.7	2199	24	AAI47130 Pyrin domain conta
27	254	4.7	2575	24	AAI70683 Human nucleotide b
28	254	4.7	2767	25	ABT16018 NOXV related polyn
29	254	4.7	3172	24	ABN99366 Human secreted pro
30	252.4	4.6	3368	24	AAI44366 Human PYRIN-3 CDNA
31	246.8	4.5	2614	25	ABV72512 Nucleotide sequenc
32	241.8	4.4	1557	24	AAI47143 Pyrin domain conta
33	237	4.4	237	22	ABA69780 Human foetal liver
34	237	4.4	237	22	ABA69780 Human genome-deriv
35	237	4.4	237	22	ABA66666 Human foetal liver
36	237	4.4	237	22	AAI17979 Human brain expres
37	237	4.4	237	22	AAI49872 Human bone marrow
38	237	4.4	237	22	AAI49872 Human bone marrow
39	237	4.4	237	22	ABA69780 Human liver single
40	226.4	4.2	469	22	ABA58208 Human genome-deriv
41	226.4	4.2	469	22	ABA58208 Human foetal liver
42	226.4	4.2	469	22	ABA27416 Probe #5882 for ge
43	226.4	4.2	469	22	AAK6371 Human brain expres
44	226.4	4.2	469	22	AAI32039 Human bone marrow
45	226.4	4.2	469	22	AAI15917 Probe #5850 for ge
					Probe #5853 used t

ALIGNMENTS

RESULT 1	
AAF83651	AAF83651 standard; CDNA; 5523 BP.
ID	AAF83651
AC	AAF83651;
XX	
DT	23-JUL-2001 (first entry)
XX	
DE	Human CARD-7 polypeptide encoding CDNA.
XX	
KW	CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;
KW	autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;
KW	antibacterial; antiviral; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	523..4812
FT	/tag= a
FT	/product= "CARD-7"
XX	
XX	MO200130813-A1.
XX	
XX	PD
XX	03-MAY-2001.
XX	
XX	27-OCT-2000; 2000WO-US29796.
XX	
XX	27-OCT-1999; 99US-0428252.
XX	
XX	(MILL-) MILLENNIUM PHARM INC.
XX	

PI Bertin J;

XX MPI, 2001-343340/36.

DR P-PSDB; AAB62571.

XX

Identifying a modulator of interaction between caspase recruitment domain (CARD)-7 and CARD-5, for treating autoimmune disorders, PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of the compound -

PS Disclosure; Fig 1A-D; 80pp; English.

XX

XX

The invention relates to identifying compounds, that modulate interaction between caspase recruitment domain (CARD)-7 and CARD-5. The method involves measuring the binding of CARD-7 and CARD-5 in the presence of the compound (an increase in the binding of CARD-7 to CARD-5 in the presence of the compound compared to the binding in the absence of the compound indicates that the compound is a modulator of CARD-7-CARD-5 interaction). Modulators of CARD-7 and CARD-8 expression or activity can be used to treat or diagnose disorders such as cancers, bacterial or viral infections, autoimmune disorders (systemic lupus erythematosus, immune-mediated glomerulonephritis or arthritis), inflammatory disorders, organ-specific autoimmunity including multiple sclerosis, Hashimoto's thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies. CARD-7 and CARD-8 are useful as modulating agents in regulating a variety of cellular processes including cell growth and cell death. The present sequence represents a cDNA encoding the human CARD-7 polypeptide.

SQ Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;

Query Match 100.0%; Score 5444; DB 22; Length 5523;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCCAGGGCTGAGAGCTGAGAGAACTGAGAGCCAGACCCGGGGCTCCACTCT 60
 DB 1 GCCCCAGGGCTGAGAGCTGAGAGAACTGAGAGCCAGACCCGGGGCTCCACTCT 60
 QY 61 GGGTTTGAAAGCCCATTCCTGCTCTGCGGCTCTCTCCACCCCACTTCTTCAGCCTT 120
 DB 61 GGGTTTGAAAGCCCATTCCTGCTCTGCGGCTCTCTCCACCCCACTTCTTCAGCCTT 120
 QY 121 GCAGCTCAAGGGTGAATCAGAGATCCAGGACCCAGAGAGGAGAAATCTAGAGAA 180
 DB 121 GCAGCTCAAGGGTGAATCAGAGATCCAGGACCCAGAGAGGAGAAATCTAGAGAA 180
 QY 181 CAGAGAGTGAAGGCTGCCACACCCCATCTCCGTCACCAACATCTCCCTCAC 240
 DB 181 CAGAGAGTGAAGGCTGCCACACCCCATCTCCGTCACCAACATCTCCCTCAC 240
 QY 241 CCTCCCTGCTGGGCTGGAACCCCATCCAGACCTTCCATAGCTGAATCTTTCAGT 300
 DB 241 CCTCCCTGCTGGGCTGGAACCCCATCCAGACCTTCCATAGCTGAATCTTTCAGT 300
 QY 301 GCTTCAGAGGCTCTGAGGCTCTCCCTCCCTGCTTTTCTACACATCCCTCTAT 360
 DB 301 GCTTCAGAGGCTCTGAGGCTCTCCCTCCCTGCTTTTCTACACATCCCTCTAT 360
 QY 361 CGGCGTCTATCTGTAGGTGCTGAGATTATATAAATGAGGTTCCGATGCTGAATA 420
 DB 361 CGGCGTCTATCTGTAGGTGCTGAGATTATATAAATGAGGTTCCGATGCTGAATA 420
 QY 421 GACGGTAAGAGCCAAAGGACAGACATGTTCTCTGCTCTGATACCTTCCAC 480
 DB 421 GACGGTAAGAGCCAAAGGACAGACATGTTCTCTGCTCTGATACCTTCCAC 480
 QY 481 CTGGGAACATCCCCAGACACCTTCTTAATCCGGAACAGATGGCTGGAGGCTGG 540
 DB 481 CTGGGAACATCCCCAGACACCTTCTTAATCCGGAACAGATGGCTGGAGGCTGG 540
 QY 541 GGGCGGCTGGGCTGTTACTTGAGGTTCTGAGAGAGAGAGCTGAAGAGTTCACTT 600
 DB 541 GGGCGGCTGGGCTGTTACTTGAGGTTCTGAGAGAGAGAGCTGAAGAGTTCACTT 600

QY 601 CTGCTGCCAATPAAGCGCACTTCAGAGAGCTTTCGGGTGAGACACCGGCTCAGCCAGAG 660
 DB 601 CTGCTGCCAATPAAGCGCACTTCAGAGAGCTTTCGGGTGAGACACCGGCTCAGCCAGAG 660
 QY 661 AAGAGAGAGGAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 661 AAGAGAGAGGAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 TGGGACCTAGCCCTCATCTGAGGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 780
 DB 721 TGGGACCTAGCCCTCATCTGAGGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 780
 QY 781 CAGAGAGGGGAGGCGCACTTCTCTCATTCCTTACAGCCCAAGTGAACCCCACTGGGG 840
 DB 781 CAGAGAGGGGAGGCGCACTTCTCTCATTCCTTACAGCCCAAGTGAACCCCACTGGGG 840
 QY 841 TCTCCAGGCAACCCACCTTCCACCCGAGGCTAATGCTGATGCTGATGCTGATGCTGATGCT 900
 DB 841 TCTCCAGGCAACCCACCTTCCACCCGAGGCTAATGCTGATGCTGATGCTGATGCTGATGCT 900
 QY 901 GGGTGCAACCCAGGCTCAGAGAGAGGTTTGAAGAGAGCTGCTGACATCTGAGAGCC 960
 DB 901 GGGTGCAACCCAGGCTCAGAGAGAGGTTTGAAGAGAGCTGCTGACATCTGAGAGCC 960
 QY 961 CGCTGAGAGAAATCTTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1020
 DB 961 CGCTGAGAGAAATCTTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1020
 QY 1021 GAGTCTCCAGGCAAGGAGTCAACCAAGCCGCTTCAATCCAGAGAGGCTGAGGAGGCTGAG 1080
 DB 1021 GAGTCTCCAGGCAAGGAGTCAACCAAGCCGCTTCAATCCAGAGAGGCTGAGGAGGCTGAG 1080
 QY 1081 GGATCCCACTCAGCCCAAGCTTCAACCAAGGCTTCAACCAAGGCTTCAACCAAGGCTTCA 1140
 DB 1081 GGATCCCACTCAGCCCAAGCTTCAACCAAGGCTTCAACCAAGGCTTCAACCAAGGCTTCA 1140
 QY 1141 CCTCTGAGTGAAGCTCAGAGAAATTTAATCAAGAAATCAAGAAATCAAGAAATCAAGAA 1200
 DB 1141 CCTCTGAGTGAAGCTCAGAGAAATTTAATCAAGAAATCAAGAAATCAAGAAATCAAGAA 1200
 QY 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
 DB 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
 QY 1261 AGCCTTACAGCCCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
 DB 1261 AGCCTTACAGCCCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
 QY 1321 TGGCCCTGGAATAATGAGGATTTTAACCAAAATTCACAGCTGCTTCTTCAAAAGA 1380
 DB 1321 TGGCCCTGGAATAATGAGGATTTTAACCAAAATTCACAGCTGCTTCTTCAAAAGA 1380
 QY 1381 CCTCAACCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 DB 1381 CCTCAACCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 QY 1441 AATCGAGAGCAATTAATGAGATCAGAGACTTAATTTGGCCAGGCTGAGTACCAAGAA 1500
 DB 1441 AATCGAGAGCAATTAATGAGATCAGAGACTTAATTTGGCCAGGCTGAGTACCAAGAA 1500
 QY 1501 CCTCGATATGATCATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
 DB 1501 CCTCGATATGATCATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
 QY 1561 GTGAAGAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
 DB 1561 GTGAAGAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
 QY 1621 TTCAAGCTGAGAGAGCTGGCCAGTCAAGAGTGTGAGTCTGCTGAGGCTCATCGAGAA 1680
 DB 1621 TTCAAGCTGAGAGAGCTGGCCAGTCAAGAGTGTGAGTCTGCTGAGGCTCATCGAGAA 1680

QY	1691	GATGGGACAGCACTCCGGGCTCCCATTAAGACAGATCCGTCTTAAAGCCAGCGGCTCTC	1740
Db	1691	GATGGGACAGCACTCCGGGCTCCCATTAAGACAGATCCGTCTTAAAGCCAGCGGCTCTC	1740
QY	1741	TTCAATCCTCAGATGGTGTATGATGAGCCAGAGATGGGTCCTTGACAGAGCCGAGTTCGAGCTC	1800
Db	1741	TTCAATCCTCAGATGGTGTATGATGAGCCAGAGATGGGTCCTTGAGAGCCGAGTTCGAGCTC	1800
QY	1801	TGTCGCACTGGAGCCAGCCACAAGCCGGGATGCACTGCTGGGCAGTTTGCTGGGAAA	1860
Db	1801	TGTCGCACTGGAGCCAGCCACAAGCCGGGATGCACTGCTGGGCAGTTTGCTGGGAAA	1860
QY	1861	ACTATACCTCCCGAGGCACTCCTTCTGATACAGGCTGGGACACAAGCTCTTGACAAACCTC	1920
Db	1861	ACTATACCTCCCGAGGCACTCCTTCTGATACAGGCTGGGACACAAGCTCTTGACAAACCTC	1920
QY	1921	ATTCCTCTTGGAGAGCAGGCACTTGGGTATAGGTCCTGGGGTCTCTGATTCAGCAGG	1980
Db	1921	ATTCCTCTTGGAGAGCAGGCACTTGGGTATAGGTCCTGGGGTCTCTGATTCAGCAGG	1980
QY	1981	AAGGAATATTTCTACAGATATTTCAACAGATGAAAGGCAAGCAATTAAGCCTTTAGGTTG	2040
Db	1981	AAGGAATATTTCTACAGATATTTCAACAGATGAAAGGCAAGCAATTAAGCCTTTAGGTTG	2040
QY	2041	GTCAAAATCAAAAGAGCTTGGGCGCTGTGTCTTGTGCGCTGGGGTCTGGCTGGCC	2100
Db	2041	GTCAAAATCAAAAGAGCTTGGGCGCTGTGTCTTGTGCGCTGGGGTCTGGCTGGCC	2100
QY	2101	TGCATTTGCCGATGACAGCAGATAGAGCGGAGGAAAAACTCACATGACCTTCCAGAGCC	2160
Db	2101	TGCATTTGCCGATGACAGCAGATAGAGCGGAGGAAAACTCACATGACCTTCCAGAGCC	2160
QY	2161	ACCAACAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAAGCTCAAGCCATTGGAGCC	2220
Db	2161	ACCAACAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAAGCTCAAGCCATTGGAGCC	2220
QY	2221	CAGCTCAGAGACCTCTGTCTCTGTGGCTGTGAGGGCATCTGGCAAAAAAGACCTTTTC	2280
Db	2221	CAGCTCAGAGACCTCTGTCTCTGTGGCTGTGAGGGCATCTGGCAAAAAAGACCTTTTC	2280
QY	2281	AGTCAGATGACCTCAGGAAGCATGGGGTATATGGGGGCATCATCTCACCTTCTTGAAG	2340
Db	2281	AGTCAGATGACCTCAGGAAGCATGGGGTATATGGGGGCATCATCTCACCTTCTTGAAG	2340
QY	2341	ATGGGTATTTCTTCAAGAGCAACCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC	2400
Db	2341	ATGGGTATTTCTTCAAGAGCAACCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC	2400
QY	2401	CAAGAGTTCCTTGCAGCAATGTCTTATGCTTGGAGATGAGAGGGGAGAGGTAAACAT	2460
Db	2401	CAAGAGTTCCTTGCAGCAATGTCTTATGCTTGGAGATGAGAGGGGAGAGGTAAACAT	2460
QY	2461	TCTAATTCGATCATATGATTTGGAAAAAGCGTTAAGAGATATGAAATACATGGCCTGTTT	2520
Db	2461	TCTAATTCGATCATATGATTTGGAAAAAGCGTTAAGAGATATGAAATACATGGCCTGTTT	2520
QY	2521	GGGCGATCAACACACGTTTCTTATTTGGGCTGTTAAGTGATGAGGGGAGAGAGAGATG	2580
Db	2521	GGGCGATCAACACACGTTTCTTATTTGGGCTGTTAAGTGATGAGGGGAGAGAGAGATG	2580
QY	2581	GAGAACATCTTTCACATGCGGCTGTCTCAGGGGAGGAACTGATGACATGGGTCCGCTCC	2640
Db	2581	GAGAACATCTTTCACATGCGGCTGTCTCAGGGGAGGAACTGATGACATGGGTCCGCTCC	2640
QY	2641	CTGCAAGCTGCTGCTGACGCCACACTCTTGGAGTCCCTCCACTGCTGTACGAGATCTGG	2700
Db	2641	CTGCAAGCTGCTGCTGACGCCACACTCTTGGAGTCCCTCCACTGCTGTGTGAGAACCTGG	2700
QY	2701	AACAAAAAGTTCCTGACACAAAGTATGGCCCATTTGCAAGAAATGGGCACTGTGTATGAA	2760
Db	2701	AACAAAAAGTTCCTGACACAAAGTATGGCCCATTTGCAAGAAATGGGCACTGTGTATGAA	2760
QY	2761	ACAGGATGAGGCTCTTAGTGTGACCTTTCGATTAATTCAGCCGCGCACGTAGAAAG	2820

Dp	2761	ACGAGCACTGGAGCTCTTAGTGTGCATTTCGACTTTAAATTCAGCCGCCACGTGAAGAAG	2820
Qy	2821	CTTCAGCTGTAATTGAGGCGCAGCAGCACAGATCAAACATGAGGCCCAACATGTAATGTCCTG	2880
Dp	2821	CTTCAGCTGTAATTGAGGCGCAGCAGCACAGATCAAACATGAGGCCCAACATGTAATGTCCTG	2880
Qy	2881	TTCAGGTGGGTCCCAGTCAACAATGCTTAATGGCAGATTCCTCTTCCGTCTCAAGTC	2940
Dp	2881	TTCAGGTGGGTCCCAGTCAACAATGCTTAATGGCAGATTCCTCTTCCGTCTCAAGTC	2940
Qy	2941	ACCAGAAACCTGAAAGAGCTGGGACCTTAATGAGGAAAACCTCGCTGAGGCCACTCTGACGTAG	3000
Dp	2941	ACCAGAAACCTGAAAGAGCTGGGACCTTAATGAGGAAAACCTCGCTGAGGCCACTCTGACGTAG	3000
Qy	3001	AGTCTTTGTATAAGACCCCTGAGACGCCCCCTCGCTGCCTCTTGAGAACCCCTGCGATTTGGCTGC	3060
Dp	3001	AGTCTTTGTATAAGACCCCTGAGACGCCCCCTCGCTGCCTCTTGAGAACCCCTGCGATTTGGCTGC	3060
Qy	3061	TGTGGCCTCAACAGTGAAGACTGCAAGAGACCTTGGCTTTGGGCTGAGAGCCAACACAGACC	3120
Dp	3061	TGTGGCCTCAACAGTGAAGACTGCAAGAGACCTTGGCTTTGGGCTGAGAGCCAACACAGACC	3120
Qy	3121	CTGACCCGAGCTGGAACCTGAGACTTCAATATGCTCTACAGGATGCTGAGGCAAAACACTTTG	3180
Dp	3121	CTGACCCGAGCTGGAACCTGAGACTTCAATATGCTCTACAGGATGCTGAGGCAAAACACTTTG	3180
Qy	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGAGCTGTCACAGCTGTGGCTC	3240
Dp	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGAGCTGTCACAGCTGTGGCTC	3240
Qy	3241	ACGTCTGACTGCTGGCCAGAGACTTGGGCTCTGTGCTTAATGTGCACGCCCCAGCCTGAAGAG	3300
Dp	3241	ACGTCTGACTGCTGGCCAGAGACTTGGGCTCTGTGCTTAATGTGCACGCCCCAGCCTGAAGAG	3300
Qy	3301	CTAGAACCTTGAGAGAGAACCAACCTGGAAGAGCTTGGCGCTGGGAGACTGCTCTGAGGGGCTC	3360
Dp	3301	CTAGAACCTTGAGAGAGAACCAACCTGGAAGAGCTTGGCGCTGGGAGACTGCTCTGAGGGGCTC	3360
Qy	3361	AGGCAATCTTCGCTGCAAACTCATACGCTGGGCTGGAGCCAGACAACCTCTGAGTATGAG	3420
Dp	3361	AGGCAATCTTCGCTGCAAACTCATACGCTGGGCTGGAGCCAGACAACCTCTGAGTATGAG	3420
Qy	3421	ATGAGGAGAGAACTGAGGGGCTCTGGAGCAGAGAAACCTCAGCTGCTCATCTTTACAGAGA	3480
Dp	3421	ATGAGGAGAGAACTGAGGGGCTCTGGAGCAGAGAAACCTCAGCTGCTCATCTTTACAGAGA	3480
Qy	3481	CGAAGAACCAAGTGTGATGATACCCCTTACTGAGGGCTGTGATACGGGAGAGATGATTAATAGC	3540
Dp	3481	CGAAGAACCAAGTGTGATGATACCCCTTACTGAGGGCTGTGATACGGGAGAGATGATTAATAGC	3540
Qy	3541	ACATCCCTCACTCAAGCGGCGAGAGACTGGATTCAGAGAGGGCGGCTTCCCAATGTTGCTCAG	3600
Dp	3541	ACATCCCTCACTCAAGCGGCGAGAGACTGGATTCAGAGAGGGCGGCTTCCCAATGTTGCTCAG	3600
Qy	3601	GCTAATCTCAAACTCCTGAGCGTGAAGCAGATCTTCCAAATGCTGATGATTCAGAGAA	3660
Dp	3601	GCTAATCTCAAACTCCTGAGCGTGAAGCAGATCTTCCAAATGCTGATGATTCAGAGAA	3660
Qy	3661	AGCTTCCCAAGAGGTAGTACCGGTGGAACCTTGTGCGCTGCTTCTCTGCTCTCAAGGG	3720
Dp	3661	AGCTTCCCAAGAGGTAGTACCGGTGGAACCTTGTGCGCTGCTTCTCTGCTCTCAAGGG	3720
Qy	3721	GACCTGCAATCGAAGCCTTTGGGGGACGAGATGACTTGGGGGCCCAACCGGGGCTGTG	3780
Dp	3721	GACCTGCAATCGAAGCCTTTGGGGGACGAGATGACTTGGGGGCCCAACCGGGGCTGTG	3780
Qy	3781	GCTACTGAGGTAGTTGACAAAGAAAAAAGACTTGTACCGAGTTCACCTTCCGTGAGCTGGC	3840
Dp	3781	GCTACTGAGGTAGTTGACAAAGAAAAAAGACTTGTACCGAGTTCACCTTCCGTGAGCTGGC	3840
Qy	3841	TCTTACCGCTGGCCCAACAGGGGTCTGCTTTGTATGAGAGAGGGGTGACCGTGTGAG	3900
Dp	3841	TCTTACCGCTGGCCCAACAGGGGTCTGCTTTGTATGAGAGAGGGGTGACCGTGTGAG	3900

```

Db      3841 TCTTACCGCTGGCCCAACAACGGGCTCTCTGATGATGAGAAAGCGGTGACCGTTGAG 3900
Qy      3901 ATTGAATTCGTGTGTGGAGCAGTCTCTGGGTGAGATCAACCAACAGCAGCTGAGT 3960
Db      3901 ATTGAATTCGTGTGTGGAGCAGTCTCTGGGTGAGATCAACCAACAGCAGCTGAGT 3960
Qy      3961 GTGGCAGGGCTCTGTGTGATCAAGGCTGAGGCTGTGAGAGCTGTGACCTTC 4020
Db      3961 GTGGCAGGGCTCTGTGTGATCAAGGCTGAGGCTGTGAGAGCTGTGACCTTC 4020
Qy      4021 CCTCATTTGTGGCTCTCAAGGGGGGCGATGTGAGACATCCCTGTTCAAAATGGCCAC 4080
Db      4021 CCTCATTTGTGGCTCTCAAGGGGGGCGATGTGAGACATCCCTGTTCAAAATGGCCAC 4080
Qy      4081 TTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCAGCGAGGTGAGAGTGCATCATAGTT 4140
Db      4081 TTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCAGCGAGGTGAGAGTGCATCATAGTT 4140
Qy      4141 CTGGAAGAACCCAGCTTCTCCCTTGGAGATCTCTGAAATGATCCATTAATGCCCTG 4200
Db      4141 CTGGAAGAACCCAGCTTCTCCCTTGGAGATCTCTGAAATGATCCATTAATGCCCTG 4200
Qy      4201 CGCTTCATTCGGTCACTCTGTGTGTGCTTTTACCAACGGGTCCATCTGAGGAAGTC 4260
Db      4201 CGCTTCATTCGGTCACTCTGTGTGTGCTTTTACCAACGGGTCCATCTGAGGAAGTC 4260
Qy      4261 ACCTTCACCTCTACCTGATCCCAAGTGAAGTCTCCATTCGGAAGAACTGAGAGCTGC 4320
Db      4261 ACCTTCACCTCTACCTGATCCCAAGTGAAGTCTCCATTCGGAAGAACTGAGAGCTGC 4320
Qy      4321 TATCGAAGCCCTGAGAGAAAGCAAGCTGTCTGAGAGTTTACGTTGGCCACTTGGAGTCA 4380
Db      4321 TATCGAAGCCCTGAGAGAAAGCAAGCTGTCTGAGAGTTTACGTTGGCCACTTGGAGTCA 4380
Qy      4381 GGGATCAGGCTGCAAGTGAAGAAAGCAAGAAAGATGAGACTCTGTGTGGAGGCGCTTGGTG 4440
Db      4381 GGGATCAGGCTGCAAGTGAAGAAAGCAAGAAAGATGAGACTCTGTGTGGAGGCGCTTGGTG 4440
Qy      4441 AAACGAGAGATCTCATGCTCTGCAACTACTCTGATCCCTCCAGCCGCAATAGCCGTACT 4500
Db      4441 AAACGAGAGATCTCATGCTCTGCAACTACTCTGATCCCTCCAGCCGCAATAGCCGTACT 4500
Qy      4501 TCACTCTGAGATGCCCCGAGTGTCTGCACTTTGTGAGCAAGTATTCGAGAGAGCTGATTA 4560
Db      4501 TCACTCTGAGATGCCCCGAGTGTCTGCACTTTGTGAGCAAGTATTCGAGAGAGCTGATTA 4560
Qy      4561 GCGCGAGTACATCGGTGAGGTTGCTTGAACAACTGATGAGCAGGTGCTGAGCCAG 4620
Db      4561 GCGCGAGTACATCGGTGAGGTTGCTTGAACAACTGATGAGCAGGTGCTGAGCCAG 4620
Qy      4621 GAGCAGTACGAGAGGCTGTGCTGAGAAACAGAGGCCAGCCAGTTCGGAAGCTGTTTC 4680
Db      4621 GAGCAGTACGAGAGGCTGTGCTGAGAAACAGAGGCCAGCCAGTTCGGAAGCTGTTTC 4680
Qy      4681 AGCTTGAAGCAGTCTCTGAGAACCGGAGTGCAGAAAGTGAAGTCTTACCAAGCCCTGAGAG 4740
Db      4681 AGCTTGAAGCAGTCTCTGAGAACCGGAGTGCAGAAAGTGAAGTCTTACCAAGCCCTGAGAG 4740
Qy      4741 ACCCATCTCAGCTCATTTATGGAACCTGAGAAAGGAGGAGCAAAAAGGAGCTCCGCGCA 4800
Db      4741 ACCCATCTCAGCTCATTTATGGAACCTGAGAAAGGAGGAGCAAAAAGGAGCTCCGCGCA 4800
Qy      4801 CTCAGAGAGTGAAGTATCAACACAGCCCTTGAAGCTTGAAGTCTGAGCTTGGAGTCAACC 4860
Db      4801 CTCAGAGAGTGAAGTATCAACACAGCCCTTGAAGCTTGAAGTCTGAGCTTGGAGTCAACC 4860
Qy      4861 TTTCTTGGGTCTCAGTTTCTTCTCTGCAAAAGATGTCATCTGTTGGCTTCCAGCA 4920
Db      4861 TTTCTTGGGTCTCAGTTTCTTCTCTGCAAAAGATGTCATCTGTTGGCTTCCAGCA 4920
Qy      4921 CTAAAGTATGAACTTGTATGATGCTTGTGCGGAGCTTATATGTGTCATGCGCAGGAGT 4980
Db      4921 CTAAAGTATGAACTTGTATGATGCTTGTGCGGAGCTTATATGTGTCATGCGCAGGAGT 4980

```

```

Qy      4981 CCACAGGGGGGCCAGTCCAGGTGCTTACAGCATCTCAAGGAATGTCCATCTGAGCT 5040
Db      4981 CCACAGGGGGGCCAGTCCAGGTGCTTACAGCATCTCAAGGAATGTCCATCTGAGCT 5040
Qy      5041 GGCAGAGCCCTCTGAGACCTCATTAAGGCTCATCTGTGTGGCCACAGCCCAAGCTTGA 5100
Db      5041 GGCAGAGCCCTCTGAGACCTCATTAAGGCTCATCTGTGTGGCCACAGCCCAAGCTTGA 5100
Qy      5101 GCGCTCCGGATCCCATCCAGGGGCAAGAGAAATGAGAGGACATGAGAACATTTGGCTC 5160
Db      5101 GCGCTCCGGATCCCATCCAGGGGCAAGAGAAATGAGAGGACATGAGAACATTTGGCTC 5160
Qy      5161 TGGCTGTGTCAAGGTGAGGCCCAAAATTTGGGTTTCAAGCTGTGAGGCCAGTGAATTC 5220
Db      5161 TGGCTGTGTCAAGGTGAGGCCCAAAATTTGGGTTTCAAGCTGTGAGGCCAGTGAATTC 5220
Qy      5221 TTTGCTTTGTACAGGAATATCTACAGAGCAAGCCAAAGTAAGTGAAGAAAGTTT 5280
Db      5221 TTTGCTTTGTACAGGAATATCTACAGAGCAAGCCAAAGTAAGTGAAGAAAGTTT 5280
Qy      5281 ATTCAAGAAATTAAGAGATATCAAGCTCTTTAGAAATTTGTCTGAGAGGCTTCCAGTT 5340
Db      5281 ATTCAAGAAATTAAGAGATATCAAGCTCTTTAGAAATTTGTCTGAGAGGCTTCCAGTT 5340
Qy      5341 TTTACAGAAAAACCCCTAATAATTTTAAATTTTAACTTAAATTTAAGATTTAAAAAAT 5400
Db      5341 TTTACAGAAAAACCCCTAATAATTTTAAATTTTAACTTAAATTTAAGATTTAAAAAAT 5400
Qy      5401 ACAAAGAAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444
Db      5401 ACAAAGAAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444

RESULT 2
ABS56030
ID ABS56030 standard; cDNA, 5523 BP.
XX
XX ABS56030;
XX
AC 10-JAN-2003 (first entry)
XX
DT
XX
DE cDNA encoding human caspase recruitment domain-7 (CARD-7).
XX
XX Human; caspase activity; caspase recruitment domain-7; CARD-7;
XX caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
XX pseudolICE; ICEBERG; cell growth; cell death; inflammation;
XX apoptosis; caspase activation; cancer; follicular lymphoma;
XX leukemia; melanoma; colon cancer; lung carcinoma; viral infection;
XX autoimmune disease; systemic lupus erythematosus; reactive arthritis;
XX human immunodeficiency virus infection; HIV infection; AIDS;
XX Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
XX myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
XX insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
XX graft rejection; allergic rhinitis; food allergy; conjunctivitis;
XX glomerular nephritis; cytostatic; virucide; immunosuppressive;
XX dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 523..4812
XX FT /*tag= a
XX FT /product= "CARD-7"
XX
XX US2002128198-A1.
XX
XX 12-SEP-2002.
XX
XX 27-NOV-2001; 2001US-0996617.
XX
XX 28-JUN-1999; 99US-0340620.
XX
XX 27-OCT-1999; 99US-0428252.

```

PR 15-AUG-2001; 2001US-0931071.
 XX (BERT/) BERTIN J.
 XX BERTIN J;
 XX WPI: 2003-028967/02.
 DR P-PSDB; ABG71633.
 XX
 PT Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
 PT CARD-7 and CARD-5 in presence of test compound, measuring their
 PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
 PT altered -
 XX
 PS Disclosure; Fig 1; 43pp; English.

XX The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudolice) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC atrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC encodes human CARD-7.

XX Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;

Query Match 100.0%; Score 5444; DB 25; Length 5523;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCCAGGAGCTGAGAGGCTGGAAGAACTGGAGGACAGAGCCGGGGCTCCACTCT 60
 DB 1 GCGCCAGGAGCTGAGAGGCTGGAAGAACTGGAGGACAGAGCCGGGGCTCCACTCT 60
 QY 61 GGGTTCTGAAGGCCCATCTCTGCTGCTGGGCTCTCCACCCCACTTTCTAGCCTT 120
 DB 61 GGGTTCTGAAGGCCCATCTCTGCTGCTGGGCTCTCCACCCCACTTTCTAGCCTT 120
 QY 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGAGCCAGAGAGGAGAAATTTGAGAAC 180
 DB 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGAGCCAGAGAGGAGAAATTTGAGAAC 180
 QY 181 CAGAGAGTGAAGGCTGGCCCAACCCCATCTCCCTCAACCACTCTCCCTCACTCAC 240
 DB 181 CAGAGAGTGAAGGCTGGCCCAACCCCATCTCCCTCAACCACTCTCCCTCACTCAC 240
 QY 241 CTTCTGCTGAGCTGAGCTGAGAGCCCACTCCAGAGCTCTCTCACTGACTTTCTTCA 300
 DB 241 CTTCTGCTGAGCTGAGCTGAGAGCCCACTCCAGAGCTCTCTCACTGACTTTCTTCA 300

QY 301 GTCTTGAGAGGCTCTGAGGCTCTCTCCCTGCTGCTTTTCTTACACACTCCCCCTAT 360
 DB 301 GTCTTGAGAGGCTCTGAGGCTCTCTCCCTGCTGCTTTTCTTACACACTCCCCCTAT 360
 QY 361 CGAGCTTATCTGATGAGTGCCTTGAGATTATATAA CTGGGTTCCGATGCTGATTA 420
 DB 361 CGAGCTTATCTGATGAGTGCCTTGAGATTATATAA CTGGGTTCCGATGCTGATTA 420
 QY 421 GACGTGAAGAGCCAGAGCAAGAGAGACAGACTGTTCTGCTGCTGATACCTTACAC 480
 DB 421 GACGTGAAGAGCCAGAGCAAGAGAGAGACAGACTGTTCTGCTGCTGATACCTTACAC 480
 QY 481 CTGGGAAATATCCCCCAAGACACCTCTTAATCCGGGAGAGAGATGCTGCGAGCTTG 540
 DB 481 CTGGGAAATATCCCCCAAGACACCTCTTAATCCGGGAGAGAGATGCTGCGAGCTTG 540
 QY 541 GCGCGCTGAGCTGTTACTTGAGATTCTGTAAGAGAGAGCTGAAGAGTTTCACTT 600
 DB 541 GCGCGCTGAGCTGTTACTTGAGATTCTGTAAGAGAGAGCTGAAGAGTTTCACTT 600
 QY 601 CTGCTGCGCAATAAAGGCGCACTCCAGAGCTCTTGGGTGAGACACCGCTCAGCCAG 660
 DB 601 CTGCTGCGCAATAAAGGCGCACTCCAGAGCTCTTGGGTGAGACACCGCTCAGCCAG 660
 QY 661 AAGACAGTGAAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 661 AAGACAGTGAAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 TGGGACTTACCTCTCACTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 DB 721 TGGGACTTACCTCTCACTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 781 CAGGAGAGGAG 840
 DB 781 CAGGAGAGGAG 840
 QY 841 TCTCCAG 900
 DB 841 TCTCCAG 900
 QY 901 GGGTGAAG 960
 DB 901 GGGTGAAG 960
 QY 961 CGCTGAG 1020
 DB 961 CGCTGAG 1020
 QY 1021 GAGTCTCAAG 1080
 DB 1021 GAGTCTCAAG 1080
 QY 1081 GAGTCTCAAG 1140
 DB 1081 GAGTCTCAAG 1140
 QY 1141 CTTCTGATGAAG 1200
 DB 1141 CTTCTGATGAAG 1200
 QY 1201 TCAAG 1260
 DB 1201 TCAAG 1260
 QY 1261 AGCTTACAG 1320
 DB 1261 AGCTTACAG 1320
 QY 1321 TGGGCTGAG 1380
 DB 1321 TGGGCTGAG 1380
 QY 1381 CTTCAAG 1440
 DB 1381 CTTCAAG 1440

```

Db 1381 CTTCAACCCAGAGCCAGATCCCTGTGCAAGAGAGCTGGCTGTATATGTGGAGAG 1440
Qy 1441 AATCAGGACATTTAATTGAGATCAGAGACTTAATTTGGCCAGGCTGGATCCCAAGAA 1500
Db 1441 AATCAGGACATTTAATTGAGATCAGAGACTTAATTTGGCCAGGCTGGATCCCAAGAA 1500
Qy 1501 CCTCGCATGTGATCTGATCTGAGAGGGGCTGCTGGAAATTTGGGAATCAACTCTGCCAGGCA 1560
Db 1501 CCTCGCATGTGATCTGATCTGAGAGGGGCTGCTGGAAATTTGGGAATCAACTCTGCCAGGCA 1560
Qy 1561 GTGAAGAAAGCCTGGGGGAGAGAGCCAGCTGTATGGGGAAACCGCTTCCAGATGCTTCTAC 1620
Db 1561 GTGAAGAAAGCCTGGGGGAGAGAGCCAGCTGTATGGGGAAACCGCTTCCAGATGCTTCTAC 1620
Qy 1621 TTCACTGAGAGAGCTGGCCAGATCCAGATCAGAGTGTGATCTGCTGAGCTCATTCGAGAAA 1680
Db 1621 TTCACTGAGAGAGCTGGCCAGATCCAGATCAGAGTGTGATCTGCTGAGCTCATTCGAGAAA 1680
Qy 1681 GATGGGACAGCCATCTCCGCTCCCATTTAGACAGATCTGTCTAGGCAAGAGCGCTGCTC 1740
Db 1681 GATGGGACAGCCATCTCCGCTCCCATTTAGACAGATCTGTCTAGGCAAGAGCGCTGCTC 1740
Qy 1741 TTCACTGATGATGTGTAGATGAGCCAGGATGGGTCTTGCAGAGCCAGCTTCTAGCTC 1800
Db 1741 TTCACTGATGATGTGTAGATGAGCCAGGATGGGTCTTGCAGAGCCAGCTTCTAGCTC 1800
Qy 1801 TGTCTGCACTGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1860
Db 1801 TGTCTGCACTGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1860
Qy 1861 ACTATATCTTCCGAGGATCTTCTCTGATCAGCGCTCGGACCAAGCTCTGCAAGAACTTC 1920
Db 1861 ACTATATCTTCCGAGGATCTTCTCTGATCAGCGCTCGGACCAAGCTCTGCAAGAACTTC 1920
Qy 1921 ATTCTTCTTGTGAGAGGAGCAAGTTGGGTAGAGGTCTGGGGTCTCTGAGTCCAGAGG 1980
Db 1921 ATTCTTCTTGTGAGAGGAGCAAGTTGGGTAGAGGTCTGGGGTCTCTGAGTCCAGAGG 1980
Qy 1981 AAGGAATATTTTACAGATATTTTACAGATGAAAGCAAGCAATTAGAGCTTTAGGTTG 2040
Db 1981 AAGGAATATTTTACAGATATTTTACAGATGAAAGCAAGCAATTAGAGCTTTAGGTTG 2040
Qy 2041 GTCAATCAAAAGAGAGCTTGGGCTCTGTGTCTTGTGTGCTTGGGTCTCTGAGTCCAG 2100
Db 2041 GTCAATCAAAAGAGAGCTTGGGCTCTGTGTCTTGTGTGCTTGGGTCTCTGAGTCCAG 2100
Qy 2101 TGCATCTGCTGATGAGAGATGAAAGGAGAGAAAGCACTGACCTGACCAAGACC 2160
Db 2101 TGCATCTGCTGATGAGAGATGAAAGGAGAGAAAGCACTGACCTGACCAAGACC 2160
Qy 2161 ACCAACAACCTCTGTCTACATTAATCTTGGCCAGGCTCTTCAAGCTCAGCCATTGGAGACC 2220
Db 2161 ACCAACAACCTCTGTCTACATTAATCTTGGCCAGGCTCTTCAAGCTCAGCCATTGGAGACC 2220
Qy 2221 CAGCTCAAGAGAGCTTGTCTCTGTGGCTGTGAGGGGATCTGGCAAAAAGAGCCCTTTTC 2280
Db 2221 CAGCTCAAGAGAGCTTGTCTCTGTGGCTGTGAGGGGATCTGGCAAAAAGAGCCCTTTTC 2280
Qy 2281 AGTCAGATGACCTCAGAGAGATGGGTTAGATGGGGCCATCATCTCCACTTCTTGAAG 2340
Db 2281 AGTCAGATGACCTCAGAGAGATGGGTTAGATGGGGCCATCATCTCCACTTCTTGAAG 2340
Qy 2341 ATGGGATATTTCAAGAGAGACCCATCCCTCTGAGCTACAGCTTCACTCACTCTGTTC 2400
Db 2341 ATGGGATATTTCAAGAGAGACCCATCCCTCTGAGCTACAGCTTCACTCACTCTGTTC 2400
Qy 2401 CAAGAGTTCTTTCAGAGCAATGTCTATGTCTTGGAGATGAGAGAGGAGAGATTAACAT 2460
Db 2401 CAAGAGTTCTTTCAGAGCAATGTCTATGTCTTGGAGATGAGAGAGGAGAGATTAACAT 2460
Qy 2461 TCTAATTCATCATGATTTGGAAAAAGAGCTAGAGCATATGAAATACATGGCCGTTT 2520
Db 2461 TCTAATTCATCATGATTTGGAAAAAGAGCTAGAGCATATGAAATACATGGCCGTTT 2520

Db 2461 TCTAATTCATCATGATTTGGAAAAAGAGCTAGAGCATATGAAATACATGGCCGTTT 2520
Qy 2521 GGGGATCAACACACAGTTTCTTATTTGGGCTGTATTAAGTATGAGGGGAGAGAGATG 2580
Db 2521 GGGGATCAACACACAGTTTCTTATTTGGGCTGTATTAAGTATGAGGGGAGAGAGATG 2580
Qy 2581 GAGAACATCTTTCACTGCCGCTGTCTCAGAGGAGAAACCTGATGCAATGGTCCCCCTC 2640
Db 2581 GAGAACATCTTTCACTGCCGCTGTCTCAGAGGAGAAACCTGATGCAATGGTCCCCCTC 2640
Qy 2641 CTGAGCTGCTGCTGAGCCACACCTCTGAGAGTCCCTCACTGCTTGAAGAGACTGG 2700
Db 2641 CTGAGCTGCTGCTGAGCCACACCTCTGAGAGTCCCTCACTGCTTGAAGAGACTGG 2700
Qy 2701 AACAAAGCTTCTGACACAAAGTATGAGCCCAATTTGAGAAATGGGATGTGTAGAA 2760
Db 2701 AACAAAGCTTCTGACACAAAGTATGAGCCCAATTTGAGAAATGGGATGTGTAGAA 2760
Qy 2761 ACAGACATGAGAGCTTTATGTGCACTTTCTGATTAATTAATCAAGCCGCAAGTGAAG 2820
Db 2761 ACAGACATGAGAGCTTTATGTGCACTTTCTGATTAATTAATCAAGCCGCAAGTGAAG 2820
Qy 2821 CTTGAGCTGATTTGAGGGGAGAGAGAGATCAATGAGAGCCCAAGATGCTAGCTG 2880
Db 2821 CTTGAGCTGATTTGAGGGGAGAGAGAGATCAATGAGAGCCCAAGATGCTAGCTG 2880
Qy 2881 TTCAAGTGGTCCAGTCAAGATGCTTATTTGGAGATTTCTTCCGCTCTCAAGTTC 2940
Db 2881 TTCAAGTGGTCCAGTCAAGATGCTTATTTGGAGATTTCTTCCGCTCTCAAGTTC 2940
Qy 2941 ACCAGAAACCTGAGAGAGCTGAGCTTAAGTGAATCTGCTGAGCACTTCAAGTGAAG 3000
Db 2941 ACCAGAAACCTGAGAGAGCTGAGCTTAAGTGAATCTGCTGAGCACTTCAAGTGAAG 3000
Qy 3001 AGCTTTTGAAGACCTGAGAGAGCCCTGAGAGCCCTGAGAGCCCTGAGAGCCCTGAGAG 3060
Db 3001 AGCTTTTGAAGACCTGAGAGAGCCCTGAGAGCCCTGAGAGCCCTGAGAGCCCTGAGAG 3060
Qy 3061 TGTGGCTCAGAGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3120
Db 3061 TGTGGCTCAGAGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3120
Qy 3121 CTGACCGAGCTGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3180
Db 3121 CTGACCGAGCTGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3180
Qy 3181 CAGAGACTGAGAGAGCCGAGCTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 3240
Db 3181 CAGAGACTGAGAGAGCCGAGCTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 3240
Qy 3241 ACGTCTGAGCTGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3300
Db 3241 ACGTCTGAGCTGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3300
Qy 3301 CTAGACCTGAGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3360
Db 3301 CTAGACCTGAGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3360
Qy 3361 AGGACCTGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3420
Db 3361 AGGACCTGCTGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3420
Qy 3421 ATGAGGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3480
Db 3421 ATGAGGAGAGAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3480
Qy 3481 CGGAAACCAATGTGATGATCCCTTATGAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3540
Db 3481 CGGAAACCAATGTGATGATCCCTTATGAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3540
Qy 3541 ACATCTCACTCAAGCCGAGAGAGCTGAGATCAGAGAGGCGGCTTCCATGTTGCTCAG 3600
Db 3541 ACATCTCACTCAAGCCGAGAGAGCTGAGATCAGAGAGGCGGCTTCCATGTTGCTCAG 3600

```

QY	3601	GCTAATCTCAAACCTCCGAGACGTGAGGAAGAATCTTCCAAATTCGTGAGATTGCAAGAGAA	3660
Db	3601	GCTAATCTCAAACCTCCGAGACGTGAGGAAGAATCTTCCAAATTCGTGAGATTGCAAGAGAA	3660
QY	3661	AGCTCCCAAGAGGATGATACGGGTGGAACCTGTGCGTGCCTTCTCGCTCTCAAGGG	3720
Db	3661	AGCTCCCAAGAGGATGATACGGGTGGAACCTGTGCGTGCCTTCTCGCTCTCAAGGG	3720
QY	3721	GACCTGCAATCGAAGCCTTTGGGGACTGACGATGACTTTGGGGCCCCACGGGGCCTGTG	3780
Db	3721	GACCTGCAATCGAAGCCTTTGGGGACTGACGATGACTTTGGGGCCCCACGGGGCCTGTG	3780
QY	3781	GCTACTGAGGATGTGACAAAGAAAAMAATTGTACCGAAGTTCATTCCTGTAGCTGAGC	3840
Db	3781	GCTACTGAGGATGTGACAAAGAAAAMAATTGTACCGAAGTTCATTCCTGTAGCTGAGC	3840
QY	3841	TCCATACGCTGAGGCCCAACAGGGGCTGTGCTTGTATGAGAGAGAGCGGTGACCGTTGAG	3900
Db	3841	TCCATACGCTGAGGCCCAACAGGGGCTGTGCTTGTATGAGAGAGAGCGGTGACCGTTGAG	3900
QY	3901	ATTGAATTTCTGTGTGTGGGACCAAGTTCCTGGGTGAGATTC AACCCACAGCAGCTGAGT	3960
Db	3901	ATTGAATTTCTGTGTGTGGGACCAAGTTCCTGGGTGAGATTC AACCCACAGCAGCTGAGT	3960
QY	3961	GTGGCAGGGCCTCTGCTGTGACATCAAGGCTGAGCTGAGAGCTGTGGAAGCTGTGCACCTC	4020
Db	3961	GTGGCAGGGCCTCTGCTGTGACATCAAGGCTGAGCTGAGAGCTGTGGAAGCTGTGCACCTC	4020
QY	4021	CCTCACCTTTGTGTGCTCCAAAGGGGGCAGTGTGAGACATCCCTGTTCCAAATGGCCAC	4080
Db	4021	CCTCACCTTTGTGTGCTCCAAAGGGGGCAGTGTGAGACATCCCTGTTCCAAATGGCCAC	4080
QY	4081	TTTAAAGAGAGGGGATGCTCTCTGGAAGACAGCAGGGGTGAGCTGCATCATATGTT	4140
Db	4081	TTTAAAGAGAGGGGATGCTCTCTGGAAGAGCAGCAGGGGTGAGCTGCATCATATGTT	4140
QY	4141	CTGAAAAACCCAGACTTCTCCCTCTGGAGTCTCTGTAATAATATCATATATGCCCTG	4200
Db	4141	CTGAAAAACCCAGACTTCTCCCTCTGGAGTCTCTGTAATAATATCATATATGCCCTG	4200
QY	4201	CGCTTCAATTCCTGACACTCTGTGTGTGTGCTTTACACCGCGCTCATCTCTGAGGAAGTC	4260
Db	4201	CGCTTCAATTCCTGACACTCTGTGTGTGTGCTTTACACCGCGCTCATCTCTGAGGAAGTC	4260
QY	4261	ACCTTCAACCTCTACCTGATCCCAAGTACGCTGCATTCGGAGAGAACTGAGCTGTGC	4320
Db	4261	ACCTTCAACCTCTACCTGATCCCAAGTACGCTGCATTCGGAGAGAACTGAGAGCTGTGC	4320
QY	4321	TATCGAAGCCTCTGAGAGAGACAGAGCTTCTCGAGATTCTACGTTGGSCCATTTGGAGTCA	4380
Db	4321	TATCGAAGCCTCTGAGAGAGACAGAGCTTCTCGAGATTCTACGTTGGSCCATTTGGAGTCA	4380
QY	4381	GGGATCAAGGCTGCAAGTGAAGACAAAGAAATGAGACTGTGTGTGGAGAGGCTTGTGTG	4440
Db	4381	GGGATCAAGGCTGCAAGTGAAGACAAAGAAATGAGACTGTGTGTGGAGAGGCTTGTGTG	4440
QY	4441	AAACCAAGAGATCATAGCTGCGCAACATATCTGATTCCTCCAGGCGCGCATAGAGCGTATCT	4500
Db	4441	AAACCAAGAGATCATAGCTGCGCAACATATCTGATTCCTCCAGGCGCGCATAGAGCGTATCT	4500
QY	4501	TCACTCTGTGATGCCCGCAGTTGCTGCACTTTGTGACCAAGTATCCAGAGCAGCTGATA	4560
Db	4501	TCACTCTGTGATGCCCGCAGTTGCTGCACTTTGTGACCAAGTATCCAGAGCAGCTGATA	4560
QY	4561	GCCCGAGTGAATGAGGTGTGTTGGAACAACTGACATGACAGGTGTGAGCAG	4620
Db	4561	GCCCGAGTGAATGAGGTGTGTTGGAACAACTGACATGACAGGTGTGAGCAG	4620
QY	4621	GAGCAGTACGAGAGGGGTGCTGAGAACAGAGGCCACAGCCAGATGCGGAAACTGTTC	4680
Db	4621	GAGCAGTACGAGAGGGGTGCTGAGAACAGAGGCCACAGCCAGATGCGGAAACTGTTC	4680

QY	4861	AGCTTGAGCCAGTCTCTGGACCGGAAAGTGCAGAAAGTGACTCTACCAAGCCCTGAAAGAG	4740
QY	4861	AGCTTGAGCCAGTCTCTGGACCGGAAAGTGCAGAAAGTGACTCTACCAAGCCCTGAAAGAG <td>4740</td>	4740
Db	4861	AGCTTGAGCCAGTCTCTGGACCGGAAAGTGCAGAAAGTGACTCTACCAAGCCCTGAAAGAG <td>4740</td>	4740
QY	4741	ACCCATCTCACTCATTTATGAACTCTGGAGAGGGGACGAAAAAGGAGACTCTGCCA <td>4800</td>	4800
Db	4741	ACCCATCTCACTCATTTATGAACTCTGGAGAGGGGACGAAAAAGGAGACTCTGCCA <td>4800</td>	4800
QY	4801	CTCAGCAGCTGAGATCAACACCGACCTTGACCTCTGAGTCTGGCTTTGGCTGACC <td>4860</td>	4860
Db	4801	CTCAGCAGCTGAGATCAACACCGACCTTGACCTCTGAGTCTGGCTTTGGCTGACC <td>4860</td>	4860
QY	4861	TTCTTTGGGTCTCAGTTTCTTTCTCTGCAAAACAAGTGGCATCTGGTTTGGCTTCCAGCA <td>4920</td>	4920
Db	4861	TTCTTTGGGTCTCAGTTTCTTTCTCTGCAAAACAAGTGGCATCTGGTTTGGCTTCCAGCA <td>4920</td>	4920
QY	4921	CTAAAGTAATGGAACCTTGATGATGCTTGCTGGGCAATATGATGATCATGCCAGGATG <td>4980</td>	4980
Db	4921	CTAAAGTAATGGAACCTTGATGATGCTTGCTGGGCAATATGATGATCATGCCAGGATG <td>4980</td>	4980
QY	4981	CCACAGGGGGCCCCAGTCCAGTGGCCCTAACAGCATCTCAGGGAAATGTCATCTGAGCT <td>5040</td>	5040
Db	4981	CCACAGGGGGCCCCAGTCCAGTGGCCCTAACAGCATCTCAGGGAAATGTCATCTGAGCT <td>5040</td>	5040
QY	5041	GGCAGAGCCCCCTGAGACTCTCAATGAGCTCATCTGCTGGCCACAGACCAAGCCCTAGA <td>5100</td>	5100
Db	5041	GGCAGAGCCCCCTGAGACTCTCAATGAGCTCATCTGCTGGCCACAGACCAAGCCCTAGA <td>5100</td>	5100
QY	5101	GGCCTCCGGATGCCATCCAGGCGCAAGAGAGAAATGAGAGGAGCATGGAACATTTGGCTC <td>5160</td>	5160
Db	5101	GGCCTCCGGATGCCATCCAGGCGCAAGAGAGAAATGAGAGGAGCATGGAACATTTGGCTC <td>5160</td>	5160
QY	5161	TGGCTGTGTCAACAGGTGAGCCCCCAAAATTTGGGGTTCAAGGTGGAGGCCACGTGATTC <td>5220</td>	5220
Db	5161	TGGCTGTGTCAACAGGTGAGCCCCCAAAATTTGGGGTTCAAGGTGGAGGCCACGTGATTC <td>5220</td>	5220
QY	5221	TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGTAAAGTGAAGAAATTT <td>5280</td>	5280
Db	5221	TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGTAAAGTGAAGAAATTT <td>5280</td>	5280
QY	5281	ATTGAGAAATTAAGAGGATATCAAGTCTTTTGAATTTGTCTAGCAGGCTTCCAGTT <td>5340</td>	5340
Db	5281	ATTGAGAAATTAAGAGGATATCAAGTCTTTTGAATTTGTCTAGCAGGCTTCCAGTT <td>5340</td>	5340
QY	5341	TTTACCGAAAAACCCCTATTAATTTTAAATTTTAACTTAATTTTAAATTAATTAATTAAT <td>5400</td>	5400
Db	5341	TTTACCGAAAAACCCCTATTAATTTTAAATTTTAACTTAATTTTAAATTAATTAATTAAT <td>5400</td>	5400
QY	5401	ACAAAAAAGAAAAATGAATTAAGAGATTAAGAGTTACTTAC <td>5444</td>	5444
Db	5401	ACAAAAAAGAAAAATGAATTAAGAGATTAAGAGTTACTTAC <td>5444</td>	5444
RESULT 3			
ABSS5497			
ID	ABSS5497	standard; cDNA; 5523 BP.	
XX	ABSS5497;		
XX	09-JAN-2003	(first entry)	
DE	CDNA encoding human caspase recruitment domain-7 (CARD-7).		
XX			
XX	Human; caspase activity; caspase recruitment domain-7; CARD-7;		
KM	caespae-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;		
KM	pseudoc1; ICEBERG; cell growth; cell death; inflammation;		
KM	apoptosis; caspase activation; cancer; follicular lymphoma;		
KM	leukemia; melanoma; colon cancer; lung cancer; viral infection;		
KM	autoimmune disease; systemic lupus erythematosus; reactive arthritis;		
KM	human immunodeficiency virus infection; HIV infection; AIDS;		
KM	Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;		
KM	myocardial infarction; stroke; inflammatory disorder; Crohn's disease;		
KM	insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;		

KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytosolic; viricide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 523..4812
 FT /tag= a
 FT /product= "CARD-7"
 PN US2002128219-A1.
 XX
 PD 12-SEP-2002.
 XX
 PF 15-AUG-2001; 2001US-0931071.
 XX
 PR 27-OCT-1999; 99US-0428252.
 XX
 PA (BERT/) BERTIN J.
 PA (ALNE/) ALNEMRI E. S.
 XX
 PI Bertin J, Alnemri ES;
 XX WPI, 2003-028968/02.
 DR P-PSDB; ABG71631.
 XX
 PT Assays for identifying compound that modulates the interaction of
 PT caspase recruitment domain-8 with a CARD-8 ligand or a compound that
 PT modulates activity of CARD-8 -
 PS Disclosure, Fig 1; 49pp; English.
 XX
 CC The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudolice) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC encodes human CARD-7.
 XX
 SQ Sequence 5523 BP; 1350 A; 1529 C; 1481 G; 1163 T; 0 other;
 Query Match 99.9%; Score 5439.2; DB 25; Length 5523;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 5441; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GCCCGAGGCGCTGAGAGTCTGAAGAACTGAGAGCCGAGAGCCGCGGCTCCACTCT 60
 Db 1 GCCCGAGGCGCTGAGAGTCTGAAGAACTGAGAGCCGAGAGCCGCGGCTCCACTCT 60

QY 61 GGGTTCTGAAGCCCATTCCTGCTCTGCGGCTCTCCACCACCTCTTCTACGCTT 120
 Db 61 GGGTTCTGAAGCCCATTCCTGCTCTGCGGCTCTCCACCACCTCTTCTACGCTT 120
 QY 121 GCAGCTCAAGGGTTGATCTCAGAGATCCAGAGCCGAGAGAGGAAATCTAGAGACA 180
 Db 121 GCAGCTCAAGGGTTGATCTCAGAGATCCAGAGCCGAGAGAGGAAATCTAGAGACA 180
 QY 181 CAGAACAGTAGAGGTGGCCACACCCCATCTCCCGGACCAACATCTCCCTCAGCTCAC 240
 Db 181 CAGAACAGTAGAGGTGGCCACACCCCATCTCCCGGACCAACATCTCCCTCAGCTCAC 240
 QY 241 CCTCCCTGCTGAGCCCTGAGACCCCATCCAGAGACTCTCCATCAGCTGACTTCTTCAG 300
 Db 241 CCTCCCTGCTGAGCCCTGAGACCCCATCCAGAGACTCTCCATCAGCTGACTTCTTCAG 300
 QY 301 GTCCTGAGAGCCCTCTGAGGCTCTCCCTCCCTGCTTCTTCTACACTCCCTCTCTAT 360
 Db 301 GTCCTGAGAGCCCTCTGAGGCTCTCCCTCCCTGCTTCTTCTACACTCCCTCTCTAT 360
 QY 361 CGGCGCTCTATCTGATGAGTCCCTGAGATTATATAAAGTGGTTCCGATCTGAATAAGA 420
 Db 361 CGGCGCTCTATCTGATGAGTCCCTGAGATTATATAAAGTGGTTCCGATCTGAATAAGA 420
 QY 421 GACGGTAGAGCCAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Db 421 GACGGTAGAGCCAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 CTGGGAACATCCCGGAG 540
 Db 481 CTGGGAACATCCCGGAG 540
 QY 541 GCGCGGCTGCTGTTACTTGAAGTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 541 GCGCGGCTGCTGTTACTTGAAGTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 601 CTGCTGCGCAATTAAGCGAGCTTCAGAGAGCTTTGGGAGAGAGAGAGAGAGAGAG 660
 Db 601 CTGCTGCGCAATTAAGCGAGCTTCAGAGAGCTTTGGGAGAGAGAGAGAGAGAGAG 660
 QY 661 AAG 720
 Db 661 AAG 720
 QY 721 TGGGACCTAGCCCTCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 Db 721 TGGGACCTAGCCCTCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 781 CAGGAAAGGGGAG 840
 Db 781 CAGGAAAGGGGAG 840
 QY 841 TCTCCAGAGCAACCCAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 841 TCTCCAGAGCAACCCAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 QY 901 GGGTGAACCCAG 960
 Db 901 GGGTGAACCCAG 960
 QY 961 CGCTGAGAGAAATCTGCTGCTCATCTCCCTCAAGAGCTTTCCAGAGCTCCCAAGCA 1020
 Db 961 CGCTGAGAGAAATCTGCTGCTCATCTCCCTCAAGAGCTTTCCAGAGCTCCCAAGCA 1020
 QY 1021 GAGTCTCAAG 1080
 Db 1021 GAGTCTCAAG 1080
 QY 1081 GGAATCCCACTCAG 1140
 Db 1081 GGAATCCCACTCAG 1140

1141 CCTGTGATGAAAAGTCAAGAAATTTTACTACAGAAATCAGAGAAAGAGAGAGAAA 1200
1141 CCTGTGATGAAAAGTCAAGAAATTTTACTACAGAAATCAGAGAAAGAGAGAGAAA 1200
1201 TCAGAGAAAGGCAAGGCCCCATGAGGAGCGGTGTAGAAACGCCCCCAAGGCGCACCC 1260
1201 TCAGAGAAAGGCAAGGCCCCATGAGGAGCGGTGTAGAAACGCCCCCAAGGCGCACCC 1260
1261 AGGCTTACAGCCCAACCAACCAAGGAGCTTCTGTGAGAGAGAGAGCTTCTGTCCACA 1320
1261 AGGCTTACAGCCCAACCAACCAAGGAGCTTCTGTGAGAGAGAGAGCTTCTGTCCACA 1320
1261 AGGCTTACAGCCCAACCAACCAAGGAGCTTCTGTGAGAGAGAGAGCTTCTGTCCACA 1320
1321 TGGCCCTGGAAAAATAGAGATTTTAAACAAAATTCACACAGCTGTACTTCTACAAAAG 1380
1321 TGGCCCTGGAAAAATAGAGATTTTAAACAAAATTCACACAGCTGTACTTCTACAAAAG 1380
1381 CCTCAACCCAGAAAGCCAGAGTCCCTGTGTCAAGAGAGAGCTGAGCTTATGTAGAGAG 1440
1381 CCTCAACCCAGAAAGCCAGAGTCCCTGTGTCAAGAGAGAGCTGAGCTTATGTAGAGAG 1440
1441 AATCGAGACATTTAATTGATGATGAGACTTATTTGGCCAGAGCTGATACCAAGAA 1500
1441 AATCGAGACATTTAATTGATGATGAGACTTATTTGGCCAGAGCTGATACCAAGAA 1500
1501 CCTCGCATAGTCACTGAGAGGGGGCTGTGAAATGGGAAAGTCAACAGTGGCAGAGAG 1560
1501 CCTCGCATAGTCACTGAGAGGGGGCTGTGAAATGGGAAAGTCAACAGTGGCAGAGAG 1560
1561 GTGAAGAAAGCCTGGGGAGAGAGCCAGCTGTATGGAGACCGCTTCCAGCATGTCTTAC 1620
1561 GTGAAGAAAGCCTGGGGAGAGAGCCAGCTGTATGGGAAACCGCTTCCAGCATGTCTTAC 1620
1621 TTCAGCTGAGAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTCGCTGAGCTCATCGAAAA 1680
1621 TTCAGCTGAGAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTCGCTGAGCTCATCGAAAA 1680
1621 TTCAGCTGAGAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTCGCTGAGCTCATCGAAAA 1680
1681 GATGGAGACCACTCGGCTCCATTAGACAGATCTGTCTAGAGGCGAGGCGTCTGCTC 1740
1681 GATGGAGACCACTCGGCTCCATTAGACAGATCTGTCTAGAGGCGAGGCGTCTGCTC 1740
1741 TTCATCTCGATGTGTAGATGAGAGAGATGGGTCTTTCAGAGAGCGAGTTCTGAGCTC 1800
1741 TTCATCTCGATGTGTAGATGAGAGAGATGGGTCTTTCAGAGAGCGAGTTCTGAGCTC 1800
1801 TGTCTGCACTGAG 1860
1801 TGTCTGCACTGAG 1860
1861 ACTATCTCTCCGAGGAGATCTCTCTGATGACAGGCTCGAGCAACAGCTCTGAGAACTTC 1920
1861 ACTATCTCTCCGAGGAGATCTCTCTGATGACAGGCTCGAGCAACAGCTCTGAGAACTTC 1920
1921 ATTCTCTCTTGGAG 1980
1921 ATTCTCTCTTGGAG 1980
1921 ATTCTCTCTTGGAG 1980
1981 AAGGAATATTTCTACAGATTTTCAAGATGAAAGCAATAGAGCTTTAGAGTTG 2040
1981 AAGGAATATTTCTACAGATTTTCAAGATGAAAGCAATAGAGCTTTAGAGTTG 2040
2041 GTCAAAATCAAAAG 2100
2041 GTCAAAATCAAAAG 2100
2101 TGCATCTGCTGATGAG 2160
2101 TGCATCTGCTGATGAG 2160
2161 ACCAGAAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCCAACTCAGGCAATGGAGAG 2220
2161 ACCAGAAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCCAACTCAGGCAATGGAGAG 2220
2221 CAGCTCAG 2280
2221 CAGCTCAG 2280

2221 CAGCTCAG 2280
2281 AGTCAGATGACCTTCAG 2340
2281 AGTCAGATGACCTTCAG 2340
2341 ATGGATATCTTCAAG 2400
2341 ATGGATATCTTCAAG 2400
2401 CAAAGATCTTTCAG 2460
2401 CAAAGATCTTTCAG 2460
2461 TCTAATTCATCATATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2461 TCTAATTCATCATATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2521 GGGGATCAACCAACAGCTTCTTATTTGGGCTGTATGATGATGAGAGAGAGAGAGAGAG 2580
2521 GGGGATCAACCAACAGCTTCTTATTTGGGCTGTATGATGATGAGAGAGAGAGAGAGAG 2580
2581 GAGAAATCTTTCAG 2640
2581 GAGAAATCTTTCAG 2640
2641 CTGAGAGCTGCTGAG 2700
2641 CTGAGAGCTGCTGAG 2700
2701 AACAAACGTTCTTCAACAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
2701 AACAAACGTTCTTCAACAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
2761 ACAGATGAGAGAGCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
2761 ACAGATGAGAGAGCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
2821 CTTAGCTGATTTGAG 2880
2821 CTTAGCTGATTTGAG 2880
2881 TTCAGGTGGTCCCAATCAAGATGCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
2881 TTCAGGTGGTCCCAATCAAGATGCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
2941 ACCAGAACTGAG 3000
2941 ACCAGAACTGAG 3000
3001 AGTCTTGTAG 3060
3001 AGTCTTGTAG 3060
3061 TGTGGCTTCAACAGCTGAG 3120
3061 TGTGGCTTCAACAGCTGAG 3120
3121 CTGACCAAGTGGAG 3180
3121 CTGACCAAGTGGAG 3180
3181 CAGAGACTGAG 3240
3181 CAGAGACTGAG 3240
3241 AGCTGATGCTGAG 3300
3241 AGCTGATGCTGAG 3300
3301 CTGAGCTGAG 3360
3301 CTGAGCTGAG 3360

Db 3301 CTAGACCTGAGAGAAACAACCTGATGACGTTGGCGTGCAGCTGCTCTGTGAGGGGCTC 3360
 QY 3361 AGGCATCTGCTCTGCAAACTCATACGCTTGGGGCTGAGCAACCAACTCTGATGATGAG 3420
 Db 3361 AGGCATCTGCTCTGCAAACTCATACGCTTGGGGCTGAGCAACCAACTCTGATGATGAG 3420
 QY 3421 ATGAGCAGAGAACTGAGGGGCTGAGCAAGAGAAACCTCAGTGTCTCATCTTCAGCAAA 3480
 Db 3421 ATGAGCAGAGAACTGAGGGGCTGAGCAAGAGAAACCTCAGTGTCTCATCTTCAGCAAA 3480
 QY 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGGCTGATACGAGAGATGATGATATAGC 3540
 Db 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGGCTGATACGAGAGATGATGATATAGC 3540
 QY 3541 ACATCTCTCACTCAAGGGGAGAGACTCGATCAGAGAGGGGCTTCCATGTTGCTCAG 3600
 Db 3541 ACATCTCTCACTCAAGGGGAGAGACTCGATCAGAGAGGGGCTTCCATGTTGCTCAG 3600
 QY 3601 GCTAATCTCAAACTCTGAGCGTGAAGCAAGATCTTCCAAATGCTGATGATGAGAA 3660
 Db 3601 GCTAATCTCAAACTCTGAGCGTGAAGCAAGATCTTCCAAATGCTGATGATGAGAA 3660
 QY 3661 AGCTCCCAAGAGTATACCGGTGAACTTGTGTGCTCTCTGCTCTCAAGG 3720
 Db 3661 AGCTCCCAAGAGTATACCGGTGAACTTGTGTGCTCTCTGCTCTCAAGG 3720
 QY 3721 GACCTGCATACGAAGCTTGTGGGACTGACGATGACTTGTGGGCTTCCAGGGGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCTTGTGGGACTGACGATGACTTGTGGGCTTCCAGGGGCTGTG 3780
 QY 3781 GCTACTGAGTATGATGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTGATGAGC 3840
 Db 3781 GCTACTGAGTATGATGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTGATGAGC 3840
 QY 3841 TCCATCCGCTGCGCCCAACAGGGTCTGTGTGTGATGAGAGAACGGGTGACCTGTGAG 3900
 Db 3841 TCCATCCGCTGCGCCCAACAGGGTCTGTGTGTGATGAGAGAACGGGTGACCTGTGAG 3900
 QY 3901 ATTGAATTTGTGTGTGAGGACCACTTCTGTGTGATGATCAACCAACGACAGCTGATG 3960
 Db 3901 ATTGAATTTGTGTGTGAGGACCACTTCTGTGTGATGATCAACCAACGACAGCTGATG 3960
 QY 3961 GTGCGAGGGGCTGTGCTGAGCATCAAGGGTGAAGCTGTGAGCTGTGAGCTGTG 4020
 Db 3961 GTGCGAGGGGCTGTGCTGAGCATCAAGGGTGAAGCTGTGAGCTGTGAGCTGTG 4020
 QY 4021 CCTCATTGTGCTCTCAAGGGGCAATGTGACATCTCTGTTCAAATGAGCCAC 4080
 Db 4021 CCTCATTGTGCTCTCAAGGGGCAATGTGACATCTCTGTTCAAATGAGCCAC 4080
 QY 4081 TTTAAAGAGAGGGGATGCTCTCTGAGAACCAAGCGGTGAGCTGATCATATAT 4140
 Db 4081 TTTAAAGAGAGGGGATGCTCTCTGAGAACCAAGCGGTGAGCTGATCATATAT 4140
 QY 4141 CTGGAAGAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAGATGATCAATATGCTG 4200
 Db 4141 CTGGAAGAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAGATGATCAATATGCTG 4200
 QY 4201 CGCTTCAATCCGCTCACTCTGTGTGCTTTACACAGCGGTCCATCTGAGGAATC 4260
 Db 4201 CGCTTCAATCCGCTCACTCTGTGTGCTTTACACAGCGGTCCATCTGAGGAATC 4260
 QY 4261 ACCTTCACTCTCACTGATCCAGTGAAGTGTCTCATTTGAGAGAACTGAGAGCTGTG 4320
 Db 4261 ACCTTCACTCTCACTGATCCAGTGAAGTGTCTCATTTGAGAGAACTGAGAGCTGTG 4320
 QY 4321 TATCGAAGCCTGAGAGAACCAAGCTGTTCTCGAGATTCTACGTTGGGCACTTGGATCA 4380
 Db 4321 TATCGAAGCCTGAGAGAACCAAGCTGTTCTCGAGATTCTACGTTGGGCACTTGGATCA 4380
 QY 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTGTGTGTGAGGCGCTTGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTGTGTGTGAGGCGCTTGTG 4440

QY 4441 AAACAGAGATCTCATGCTTGCAACTACTGTGATCCCTCCAGCCCGCATACCCGATCT 4500
 Db 4441 AAACAGAGATCTCATGCTTGCAACTACTGTGATCCCTCCAGCCCGCATACCCGATCT 4500
 QY 4501 TCACTCTGAGAGCCCGGAGTTGCTGACCTTTGTGAGCAATATTCAGAGAGAGAGAG 4560
 Db 4501 TCACTCTGAGAGCCCGGAGTTGCTGACCTTTGTGAGCAATATTCAGAGAGAGAGAG 4560
 QY 4561 GCCCAGTACATCGGTGAGGTTGCTGTGACCAACTGATGACAGGTGTGAGCCAG 4620
 Db 4561 GCCCAGTACATCGGTGAGGTTGCTGTGACCAACTGATGACAGGTGTGAGCCAG 4620
 QY 4621 GAGCAGTACGAGAGGTTGCTGTGAGCAAGAGCCCAAGAGCTGTGCTGCTG 4680
 Db 4621 GAGCAGTACGAGAGGTTGCTGTGAGCAAGAGCCCAAGAGCTGTGCTGCTG 4680
 QY 4681 AGCTTGAAGCAGTCTCGGAGCCGGAAGTGAAGTGAAGTCTTACCAAGCCCTGAAGAG 4740
 Db 4681 AGCTTGAAGCAGTCTCGGAGCCGGAAGTGAAGTGAAGTCTTACCAAGCCCTGAAGAG 4740
 QY 4741 ACCCATCTCACTCATTTATGAACTGTGAGAGAGGAGCAAGAAAGGAGCTCTGCCA 4800
 Db 4741 ACCCATCTCACTCATTTATGAACTGTGAGAGAGGAGCAAGAAAGGAGCTCTGCCA 4800
 QY 4801 CTGAGAGCTGAAGTATCAACACAGCCCTGACCTTGAATCTGTGCTGTGAGCTG 4860
 Db 4801 CTGAGAGCTGAAGTATCAACACAGCCCTGACCTTGAATCTGTGCTGTGAGCTG 4860
 QY 4861 TTTCTTGGGCTCAGTTCTTCTGCAAGAAAGAGTGCATCTGCTTGGCTTCCAGCA 4920
 Db 4861 TTTCTTGGGCTCAGTTCTTCTGCAAGAAAGAGTGCATCTGCTTGGCTTCCAGCA 4920
 QY 4921 CTAAAGTATGAACTTGTGATGATGATGCTTGTGCTGAGCATATATGTCTCATGCAAGGATG 4980
 Db 4921 CTAAAGTATGAACTTGTGATGATGATGCTTGTGCTGAGCATATATGTCTCATGCAAGGATG 4980
 QY 4981 CCAAGGGGGGCTCAGTCCAGGTGCTTACAGCATCTCAGGGAATGTCTCATGAGCT 5040
 Db 4981 CCAAGGGGGGCTCAGTCCAGGTGCTTACAGCATCTCAGGGAATGTCTCATGAGCT 5040
 QY 5041 GGCAAGAACCCCTGAGACCTCTAGAGCTCATCTGAGGAGCCAGACAGCCAGCTGAG 5100
 Db 5041 GGCAAGAACCCCTGAGACCTCTAGAGCTCATCTGAGGAGCCAGACAGCCAGCTGAG 5100
 QY 5101 GCCCTCCGATCCATCCAGGCGCAAGAGAGAAATGAGAGGAGCATGAACTTTCCTC 5160
 Db 5101 GCCCTCCGATCCATCCAGGCGCAAGAGAGAAATGAGAGGAGCATGAACTTTCCTC 5160
 QY 5161 TGGCTGTGTCAAGGTTGAGCCCAAAATTTGGGGTTCAAGCTGTGAGAGGCAAGTTC 5220
 Db 5161 TGGCTGTGTCAAGGTTGAGCCCAAAATTTGGGGTTCAAGCTGTGAGAGGCAAGTTC 5220
 QY 5221 TTGGCTTTTACAGGAAGTCTCAAGAGCAAGCCCAAGATTAAGTGAAGGAAGTT 5280
 Db 5221 TTGGCTTTTACAGGAAGTCTCAAGAGCAAGCCCAAGATTAAGTGAAGGAAGTT 5280
 QY 5281 ATTCAAGAAATTAAGAGATGATCAAGCTCTTTTGAATTTGTCTAGCAGGCTTTCAGTT 5340
 Db 5281 ATTCAAGAAATTAAGAGATGATCAAGCTCTTTTGAATTTGTCTAGCAGGCTTTCAGTT 5340
 QY 5341 TTTACCAAGAAACCCCTATTAATTTAAATTTTAACTTAAATTTAAAGATTTAAAGAAAT 5400
 Db 5341 TTTACCAAGAAACCCCTATTAATTTAAATTTTAACTTAAATTTAAAGATTTAAAGAAAT 5400
 QY 5401 AAAAAAAAAAATAAATAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444
 Db 5401 AAAAAAAAAAATAAATAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444

RESULT 4
 AAL47127
 ID AAL47127 standard; DNA; 5100 BP.

XX AAL47127;
 XX 20-AUG-2002 (first entry)
 XX Pyrin domain containing protein NALP1-hs coding sequence.
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KM antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KM neuroprotective; antiallergic; antirheumatic; antiaesthetic;
 KM nephroprotective; osteoprotective; neotropic; intracellular signal transduction;
 KM inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KM arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KM osteoarthritis; glomerulonephritis; gene; ds.
 XX Unidentified.
 OS WO200240668-A2.
 XX 23-MAY-2002.
 PD 30-OCT-2001; 2001WO-EP12545.
 XX 15-NOV-2000; 2000DE-1056687.
 XX 30-NOV-2000; 2000DE-1059595.
 XX (APOT-) APOTECHE RES & DEV LTD.
 XX Techopp J, Martinson F;
 PI WPI; 2002-427093/45.
 DR P-PSDB; AAO17855.
 XX New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies -
 XX Claim 5; Fig 1; 116bp; German.
 XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention.
 XX
 SQ Sequence 5100 BP; 1251 A; 1376 C; 1404 G; 1069 T; 0 other;
 Query Match 89.6%; Score 4879.2; DB 24; Length 5100;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 4896; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 516 GACAGAGATGGCTGGCGAGGCTGGGGCCGCTGCTTACTTGAAGTTCTGAAGAA 575
 DB 153 GAACAAATGCTGCGGAGGCTGGGGCCGCTGCTTACTTGAAGTTCTGAAGAA 212
 QY 576 GGAGGAGCTGAAGAGTTCCAGCTTCTGCTGCGCAATTAAGCGACCCGAGAGCTTTC 635
 DB 213 GGAAGAGCTGAAGAGTTCCAGCTTCTGCTGCGCAATTAAGCGACCTCAGAGAGCTTTC 272
 QY 636 GGGTGAAGCAACCCGCTGAGCCAGAGAAAGAGTGGCATGAGAGTGGCTTGTGTA 695
 DB 273 GGGTGAAGCAACCCGCTGAGCCAGAGAAAGAGTGGCATGAGAGTGGCTTGTGTA 332
 QY 696 GGGTGAAGTGGGAGAGAGCGGGGCTGGGACCTTGAATCTTGAATCTTGAATCTTGA 755
 DB 333 GGGTGAAGTGGGAGAGAGCGGGGCTGGGACCTTGAATCTTGAATCTTGAATCTTGA 392
 QY 756 GCTGAGGCTGATGCGGCGCAAGCCAGAGAGGGGAGGAGGAGGAGGAGGAGGAGGAG 815
 |||||||

DB 393 GCTGAGGTCACTGTGGGCCCAAGCCCAAGAGAGGGGCAAGGCACTTCCCTCATTTCCCTTA 452
 QY 816 CAGCCCAAGTGAACCCCACTGGGGGTCTCCAGCCCAACCCCACTCCAGCGAGTCTAAT 875
 DB 453 CAGCCCAAGTGAACCCCACTGGGGGTCTCCAGCCCAACCCCACTCCAGCGAGTCTAAT 512
 QY 876 GGCCTGATCCAGAAATGCGGGGAGGAGTCCAGCGGCTCAGAGAGAGAGGAGGTTTGGAG 935
 DB 513 GGCCTGATCCAGAAATGCGGGGAGGAGTCCAGCGGCTCAGAGAGAGAGGAGGTTTGGAG 572
 QY 936 ACAGTGCCTGACACATCTTGAAGCGCGCTGAGAGAGAAATCTTGCCTCACTCTTACCA 995
 DB 573 ACAGTGCCTGACACATCTTGAAGCGCGCTGAGAGAGAAATCTTGCCTCACTCTTACCA 632
 QY 996 AGCTTTTCAAGCTCCCGACAGCATGATGTCCAGCCAGAGAGTCAACCGCCCGAC 1055
 DB 633 AGCTTTTCAAGCTCCCGACAGCATGATGTCCAGCCAGAGAGTCAACCGCCCGAC 692
 QY 1056 ATCCACAGCAGTGTGGGGAGTGGGGGATGCCACCTCAGCCAGGCTTACCCACAGAGA 1115
 DB 693 ATCCACAGCAGTGTGGGGAGTGGGGGATGCCACCTCAGCCAGGCTTACCCACAGAGA 752
 QY 1116 GCAGAGGCTCTTGGAGACCAATGAGCTCTGATGAAACGTCAAGAAATTTACTACAGAGA 1175
 DB 753 GCAGAGGCTCTTGGAGACCAATGAGCTCTGATGAAACGTCAAGAAATTTACTACAGAGA 812
 QY 1176 AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1235
 DB 813 AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 872
 QY 1236 AGGAG 1295
 DB 873 AGGAG 932
 QY 1296 TGTGAG 1355
 DB 933 TGTGAG 992
 QY 1356 CACACAGCTGCTACTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1415
 DB 993 CACACAGCTGCTACTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052
 QY 1416 AACCTGGCTGATTAATGAG 1475
 DB 1053 AACCTGGCTGATTAATGAG 1112
 QY 1476 TGGCCAG 1535
 DB 1113 TGGCCAG 1172
 QY 1536 TGGAGAGTCAACCTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1595
 DB 1173 TGGAGAGTCAACCTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
 QY 1596 GGAACGCTTCAACAGTCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1655
 DB 1233 GGAACGCTTCAACAGTCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1292
 QY 1656 GAGTCTGAGTGAAGTCAAG 1715
 DB 1293 GAGTCTGAGTGAAGTCAAG 1352
 QY 1716 CTTGTCTTGAAGCAG 1775
 DB 1353 CTTGTCTTGAAGCAG 1412
 QY 1776 CTTGACAG 1835
 DB 1413 CTTGACAG 1472
 QY 1836 ACTGCTGGAG 1895
 DB 1473 ACTGCTGGAG 1532

QY 1896 TCGGACCAAGCTCTGCAAACTCATCTCTTTGAGACAGGCACTGGGTAGAGGT 1955
DB 1533 TCGGACCAAGCTCTGCAAACTCATCTCTTTGAGACAGGCACTGGGTAGAGGT 1592
QY 1956 CCGGGGTTCTGAGCCGAGGAGAAATTTCTACAGATATTTCACAGATGAAG 2015
DB 1593 CCGGGGTTCTGAGCCGAGGAGAAATTTCTACAGATATTTCACAGATGAAG 1652
QY 2016 GCAAGCAATTAGACCTTTAGTTGTGCAATCAAAACAAGACCTGGGCTGTGTCT 2075
DB 1653 GCAAGCAATTAGACCTTTAGTTGTGCAATCAAAACAAGACCTGGGCTGTGTCT 1712
QY 2076 TGTGCTCTGGGTGTCTGTGCTGCTGCACTTGTGATGACAGCAATGAAGCGAAGA 2135
DB 1713 TGTGCTCTGGGTGTCTGTGCTGCTGCACTTGTGATGACAGCAATGAAGCGAAGA 1772
QY 2136 AAAAACAACGATGCTTCAAGACCAACCAACCTGTGCTCACTTACCTGGCCAGAG 2195
DB 1773 AAAAACAACGATGCTTCAAGACCAACCAACCTGTGCTCACTTACCTGGCCAGAG 1832
QY 2196 TCTCAAGCTCAAGCAATGGGACCCAGCTCAGAGACCTGTGCTCTGTGCTGAGAG 2255
DB 1833 TCTCAAGCTCAAGCAATGGGACCCAGCTCAGAGACCTGTGCTCTGTGCTGAGAG 1892
QY 2256 CATCTGGCAAAAAAGACCTTTTCACTCAATGACCTCAGAGAGCATGGTTAGATG 2315
DB 1893 CATCTGGCAAAAAAGACCTTTTCACTCAATGACCTCAGAGAGCATGGTTAGATG 1952
QY 2316 GGGCATCATCTCCACCTCTGTAAGATGGGTATTCTTCAAGAGACCCCACTCCCTGAG 2375
DB 1953 GGGCATCATCTCCACCTCTGTAAGATGGGTATTCTTCAAGAGACCCCACTCCCTGAG 2012
QY 2376 CTACAGCTTCACTCACTCTGTTCAGAGATCTTTCAGAGATGCTTATGCTTGA 2435
DB 2013 CTACAGCTTCACTCACTCTGTTCAGAGATCTTTCAGAGATGCTTATGCTTGA 2072
QY 2436 GGATGGAAGGGAGAGATTAATTTCAATTCATCATATGATTGGAAAAAGCGTTAG 2495
DB 2073 GGATGGAAGGGAGAGATTAATTTCAATTCATCATATGATTGGAAAAAGCGTTAG 2132
QY 2496 AGCATATGAAATCATAGGCTGTGTTGGGGCATCAACACAGCTTCTTATGGGCTGT 2555
DB 2133 AGCATATGAAATCATAGGCTGTGTTGGGGCATCAACACAGCTTCTTATGGGCTGT 2192
QY 2556 AAGTGAAGAGGGAGAGAGATGAGAACATCTTTCATCTGCGGCTGTCTCAGGGAG 2615
DB 2193 AAGTGAAGAGGGAGAGAGATGAGAACATCTTTCATCTGCGGCTGTCTCAGGGAG 2252
QY 2616 GAACCTGATGCAATGGGTCCCGTCCCTGCAAGCTGCTGACAGCCACATCTTGAAGTC 2675
DB 2253 GAACCTGATGCAATGGGTCCCGTCCCTGCAAGCTGCTGACAGCCACATCTTGAAGTC 2112
QY 2676 CCTCACTGCTTGAAGAGCTCGGAAACAAAAGCTTCTGACACAAGTATGGCCATT 2735
DB 2313 CCTCACTGCTTGAAGAGCTCGGAAACAAAAGCTTCTGACACAAGTATGGCCATT 2372
QY 2736 CGAAGAAATGGCATGTGTAGAAAACAAGATGAGCTTTAGTGTGCACTTTGCAAT 2795
DB 2373 CGAAGAAATGGCATGTGTAGAAAACAAGATGAGCTTTAGTGTGCACTTTGCAAT 2432
QY 2796 TAAATTCAGCCGCAAGTGAAGAAAGCTTCAAGCTGATTTAGGGGAGGACAGATCAAC 2855
DB 2433 TAAATTCAGCCGCAAGTGAAGAAAGCTTCAAGCTGATTTAGGGGAGGACAGATCAAC 2492
QY 2856 ATGAGACCCCAAGATGATGCTGTTCAGGTGGTCCAGTCAAGATGCTATTGGCA 2915
DB 2493 ATGAGACCCCAAGATGATGCTGTTCAGGTGGTCCAGTCAAGATGCTATTGGCA 2552
QY 2916 GATTTCTTCTCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGACCTTAAGTGA 2975
DB 2553 GATTTCTTCTCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGACCTTAAGTGA 2612

QY 2976 CTGCTGAGCACTGTGCAATGATGATGATTTGTAAAGACCTTGAGACGCGCTCGCTCT 3035
DB 2613 CTGCTGAGCACTGTGCAATGATGATGATTTGTAAAGACCTTGAGACGCGCTCGCTCT 2672
QY 3036 CCGAGAGACCTCGGAGTTGGCTGAGTGGCTCAAGCTGAGAGACTGAGAGACCTTGC 3095
DB 2673 CCGAGAGACCTCGGAGTTGGCTGAGTGGCTCAAGCTGAGAGACTGAGAGACCTTGC 2732
QY 3096 CTTTGGGCTGAGAGCCCAACAGACCTTGAACGAGCTGAGCCTGAGCTTCAATGTGCTAC 3155
DB 2733 CTTTGGGCTGAGAGCCCAACAGACCTTGAACGAGCTGAGCCTGAGCTTCAATGTGCTAC 2792
QY 3156 GATGCTGAGAGCCCAACACTTTGCAAGATGAGACAGCCGAGTGTGAACCTACAGG 3215
DB 2793 GATGCTGAGAGCCCAACACTTTGCAAGATGAGACAGCCGAGTGTGAACCTACAGG 2852
QY 3216 ACTGAGCTGTGATGATGAGTGTGCTCAAGCTGATGATGAGAGCTGAGCTGTGTGT 3275
DB 2853 ACTGAGCTGTGATGATGAGTGTGCTCAAGCTGATGATGAGAGCTGAGCTGTGTGT 2912
QY 3276 TAGTGCAGCCCAAGCTGAGAGAGCTGAGAGCTGAGAGCAACCTGATGAGCTTGG 3335
DB 2913 TAGTGCAGCCCAAGCTGAGAGAGCTGAGAGCTGAGAGCAACCTGATGAGCTTGG 2972
QY 3336 CGTGGAGCTGTGTGAGAGGCTCAAGGATCTGTGCTGCAACCTATGCTGGGCT 3395
DB 2973 CGTGGAGCTGTGTGAGAGGCTCAAGGATCTGTGCTGCAACCTATGCTGGGCT 3032
QY 3396 GGAACGAGCAACTGAGTGAATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3455
DB 3033 GGAACGAGCAACTGAGTGAATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3092
QY 3456 ACCTGAGCTGTGATCTTTCAGAGACGAGAAACCAAGTGTATGATGATGATGATGAT 3515
DB 3093 ACCTGAGCTGTGATCTTTCAGAGACGAGAAACCAAGTGTATGATGATGATGATGATGAT 3152
QY 3516 GGAATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3575
DB 3153 GGAATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3212
QY 3576 GAGGAGGCTTCCCATGTTGTCTCAGGCTATCTCAAACTCTGAGAGTGAAGATCTT 3635
DB 3213 GAGGAGGCTTCCCATGTTGTCTCAGGCTATCTCAAACTCTGAGAGTGAAGATCTT 3272
QY 3636 CCCAATGCTGATGATGAG 3695
DB 3273 CCCAATGCTGATGATGAG 3332
QY 3696 CGTGCTTCTCTGCTCTCAAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3755
DB 3333 CGTGCTTCTCTGCTCTCTCAAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3392
QY 3756 CTTTGGGAGCCCAAGGGGCTGTGAGCTACTAGAGTATGTTGCAAAAGAAAGACTTGA 3815
DB 3393 CTTTGGGAGCCCAAGGGGCTGTGAGCTACTAGAGTATGTTGCAAAAGAAAGACTTGA 3452
QY 3816 CCGAGTTCACTTCCCTGTAGCTGAGCTCTTACCGCTGAGCCCAACAGGGCTCTGCTTGT 3875
DB 3453 CCGAGTTCACTTCCCTGTAGCTGAGCTCTTACCGCTGAGCCCAACAGGGCTCTGCTTGT 3512
QY 3876 GATGAG 3935
DB 3513 GATGAG 3572
QY 3936 GATCAACCAACAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3995
DB 3573 GATCAACCAACAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3632
QY 3996 TGAAGCTGTGAG 4055
DB 3633 TGAAGCTGTGAG 3692
QY 4056 CACATCCCTGTTCAAAATGGCCCACTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4115

Db	3693	CACATCCCTGTTCCAAAGTGGCCCACTTTAAAGAGAGGAGATGCTCCTGGAGAAACGACG	3752
Qy	4116	CAGGTGGAGCTGCATCATATAGTTCTGGAAAACCCAGCTTCTCCCTTGGAGTCT	4175
Db	3753	CAGGTGGAGCTGCATCATATAGTTCTGGAAAACCCAGCTTCTCCCTTGGAGTCT	3812
Qy	4176	CCTGAAAATGATTCATTAATGCCCCGTGCTCATTTCCCGTCACTCTGAGGTGCTTTA	4235
Db	3813	CCTGAAAATGATTCATTAATGCCCCGTGCTCATTTCCCGTCACTCTGAGGTGCTTTA	3872
Qy	4236	CCACCGCGTCATCCTGAGAAAGTCACTTCCACTGATCCCAAGTGACTGCTC	4295
Db	3873	CCACCGCGTCATCCTGAGAAAGTCACTTCCACTGATCCCAAGTGACTGCTC	3932
Qy	4296	CATTTCGAGGAACCTGAGCTCTGCTATCGAAAGCCCTGGAGAAACACAGCTGTTCTGG	4355
Db	3933	CATTTCGAGGAACCTGAGCTCTGCTATCGAAAGCCCTGGAGAAACACAGCTGTTCTGG	3992
Qy	4356	GTTCTACGTTGGSCACTTGGGATCAGAGGATCAGCGCTGAAATGAAAGCAAGAAAGATGA	4415
Db	3993	GTTCTACGTTGGSCACTTGGGATCAGAGGATCAGCGCTGAAATGAAAGCAAGAAAGATGA	4052
Qy	4416	GACTCTGCTGTTGGAGAGCCTTGTGTGAACACAGAGATCTCATGCTCGCACTACCTGAT	4475
Db	4053	GACTCTGCTGTTGGAGAGCCTTGTGTGAACACAGAGATCTCATGCTCGCACTACCTGAT	4112
Qy	4476	CCCTTCAGACCCGATAGCCCGTACCTTCACTCTGGATGCCCGCAGTTGCTGCACTTGT	4535
Db	4113	CCCTTCAGACCCGATAGCCCGTACCTTCACTCTGGATGCCCGCAGTTGCTGCACTTGT	4172
Qy	4536	GGACCAATATGAGAGACAGCTGATAGCCCGAGTGCATCGGTGAGAGTTGTCTTGGACA	4595
Db	4173	GGACCAATATGAGAGACAGCTGATAGCCCGAGTGCATCGGTGAGAGTTGTCTTGGACA	4232
Qy	4596	ACTGATATGACAGGAGCTGAGCCAGAGACAGTACGAGAGGGTGGCTGAGAACACGAG	4655
Db	4233	ACTGATATGAGAGAGGTGCTGAGCCAGAGACAGTACGAGAGGGTGGCTGAGAACACGAG	4292
Qy	4656	GCCGACGCGAATGCGGAACTGTTCACTTGAAGCAGTCTTGGAGCCGGAAGTGCACAA	4715
Db	4293	GCCGACGCGAATGCGGAACTGTTCACTTGAAGCAGTCTTGGAGCCGGAAGTGCACAA	4352
Qy	4716	TGACATCTACCAAGCCCTGAAAGAGACCCATCTCTCACTTATGAACTCTGGAGAA	4775
Db	4353	TGACATCTACCAAGCCCTGAAAGAGACCCATCTCTCACTTATGAACTCTGGAGAA	4412
Qy	4776	GGGACAGCAAAAAGGAGCTCTGCACTCAGACGCTGAAAGTATGAACACACAGCCCTTGA	4835
Db	4413	GGGACAGCAAAAAGGAGCTCTGCACTCAGACGCTGAAAGTATGAACACACAGCCCTTGA	4472
Qy	4836	CTTGAGTCTGAGCTTTGGCTGACCTTCTTGGGCTCAGATTTCTTCTCTGCAAAACAG	4895
Db	4473	CTTGAGTCTGAGCTTTGGCTGACCTTCTTGGGCTCAGATTTCTTCTCTGCAAAACAG	4532
Qy	4896	TTGCAATCTGATTTGCTTCCAGCATTAAGATATGAACTTTGATGATGCTTGTCTGG	4955
Db	4533	TTGCAATCTGATTTGCTTCCAGCATTAAGATATGAACTTTGATGATGCTTGTCTGG	4592
Qy	4956	GCATTATGTGTCAATGCCAGGGATGCCACAGGGGGCCCAAGTCCAAGGTGCCTTAACAG	5015
Db	4593	GCATTATGTGTCAATGCCAGGGATGCCACAGGGGGCCCAAGTCCAAGGTGCCTTAACAG	4652
Qy	5016	TCTCAGGGGAATGTCATCTGGAGCTGGCAAGACCCCTGACAGCCCATAGAGCCCATCT	5075
Db	4653	TCTCAGGGGAATGTCATCTGGAGCTGGCAAGACCCCTGACAGCCCATAGAGCCCATCT	4712
Qy	5076	GGTGGCCACAGACAGCAAGCCTAGAGCCCTCGATCCCATTCAGAGCGCAAGAGAGATA	5135
Db	4713	GGTGGCCACAGACAGCAAGCCTAGAGCCCTCGAGTCCCATTCAGAGCGCAAGAGAGATA	4772
Qy	5136	GGAGGACATGGAACCATTTGCTTGGCTGTGTCAAGGGTAGCCCAAAATTGGGGT	5195

Db	4773	GGAGGACATCGAACAATTGGCTTCGTGCTGTGTGCACAGG9TGAGCCCAAAATTTGGGGT	4832
Oy	5196	TCAGCGTGGGAAGGCCACGCTGGATTCTTGCGCTTTGTACAGGAAGATCTTACAAAGCAAACC	5255
Db	4833	TCAGCGTGGGAAGGCCACGCTGGATTCTTGCGCTTTGTACAGGAAGATCTTACAAAGCAAAGCC	4892
Oy	5256	AACAGACTTAAGTGGAGAAGAGTTTATTCAGAAAATAAAGAGATACAGAGCTCTTTAG	5315
Db	4893	AACAGACTTAAGTGGAGAAGAGTTTATTCAGAAAATAAAGAGATATCAGTGCCTTTTAG	4952
Oy	5316	AAATTGCTCAGCAGGCTTTCCAGTTTTTACAGAAAACCCCTATAAATTAAAAATTTTTT	5375
Db	4953	AAATTGCTCAGCAGACTTCCAGTTTTTACAGAAAACCCCTATAAATTAAAAATTTTTT	5012
Oy	5376	ACTTAATAATTTAAGATTTAAAAAANTACAAAAAGAAAAATGAATAATTAAGCAATTAAGA	5435
Db	5013	ACTTAATAATTTAAGATTTAAAAAANTACAAAAAGAAAAATGAATAATTAAGCAATTAAGA	5072
Oy	5436	GTTA 5439	
Db	5073	GTTA 5076	
<hr/>			
RESULT 5			
ID	ADD12951		
XX	ADD12951 standard; cDNA; 5122 BP.		
XX	ADD12951;		
XX	16-Oct-2001 (first entry)		
DE	Human G-protein coupled receptor-8 (GCRC-8) cDNA.		
XX			
KW	Human; G-protein coupled receptor-8; GCRC-8; cytosolic; hepatotropic;		
KW	vitronic; antiinflammatory; anticonvulsant; analgesic; neuroprotective;		
KW	noctropic; cerebroprotective; hypotensive; tranquilizer; vulnerary;		
KW	ophthalmological; cell proliferative disorder; actinic keratosis;		
KW	aneurismic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;		
KW	psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease;		
KW	Huntington's disease; Parkinson's disease; cardiovascular disorder;		
KW	epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia;		
KW	anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder;		
KW	Addison's disease; Crohn's disease; acquired immune deficiency syndrome;		
KW	AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity;		
KW	osteoporosis; transgenic animal; gene therapy; ss.		
XX			
OS	Homo sapiens.		
TH			
Key	Location/Qualifiers		
FT	489..4910		
FT	/tag= a		
FT	/product= "Human GCRC-8 protein"		
XX			
PN	W0200157085-A2.		
XX			
PD	09-AUG-2001.		
PF	01-FEB-2001; 2001WO-US03455.		
XX			
PR	02-FEB-2000; 2000US-0180093.		
PR	11-FEB-2000; 2000US-0182045.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Baughn MR, Au-Young J, Yue H;		
DR	WPI; 2001-488869/53.		
DR	P-PSDB; AAB06758.		
PT	Novel isolated human G-protein coupled receptor useful for diagnosing,		
PT	preventing and treating cell proliferative, neurological,		
PT	cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic		
PT	disorders -		

XX Claim 5; Page 131-132; 138bp; English.

XX The present sequence is human G-protein coupled receptor-8 (GRCR-8)

CC cDNA. The present invention relates to GRCR protein and nucleic acids

CC encoding them. GRCR protein, its agonist or antagonist are useful for

CC treating diseases or conditions associated with decreased expression

CC or overexpression of functional GRCR in a patient, where the disorder

CC is selected from cell proliferative disorders such as actinic keratosis,

CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and

CC cancer, neurological disorders such as epilepsy, stroke, Alzheimer's

CC disease, Huntington's disease, Parkinson's disease, cardiovascular

CC disorders such as hypertension, vasculitis, varicose veins, gastro-

CC intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,

CC pancreatitis, autoimmune/inflammatory disorders such as acquired

CC immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,

CC leish, viral, bacterial, fungal, parasitic, protozoal, helminthic

CC infections, trauma and metabolic disorders such as diabetes, obesity,

CC osteoporosis. GRCR proteins and their cDNAs are used to assess the

CC effects of exogenous compounds on the expression of GRCR sequences.

CC GRCR cDNA is useful to create knock in humanised animals (pigs) or

CC transgenic animals (mice or rats) to model human disease, for

CC therapeutic or diagnostic purposes, for somatic or germline gene

CC therapy, to generate hybridisation probes useful in mapping the

CC naturally occurring genomic sequence, and in molecular biological

CC techniques.

XX

XX Sequence 5122 BP; 1184 A; 1460 C; 1401 G; 1077 T; 0 other;

XX

Query Match 87.5%; Score 4765.6; DB 22; Length 5122;

Best Local Similarity 96.9%; Pred. No. 0;

Matches 4933; Conservative 0; Mismatches 24; Indels 133; Gaps 2;

QY 35 GAGCCAGACGACCCGGGCTCCACTGGTTCGAAAGCCATTCCTGCTCGGCTC 94

DB 1 GAGCCAGACGACCCGGGCTCCACTGGTTCGAAAGCCATTCCTGCTCGGCTC 60

QY 95 CTCCACACCCCACTCTTCTAGGCTTCAGAGCTCAAGGTTGATCTCAGAGTCCAGACC 154

DB 61 CTCCACACCCCACTCTTCTAGGCTTCAGAGCTCAAGGTTGATCTCAGAGTCCAGACC 120

QY 155 CAGAGAGGGAAGATCTGAGAAACAGAAAGTGAGCGTTGCCACACCCCATCTCCC 214

DB 121 CAGAGAGGGAAGATCTGAGAAACAGAAAGTGAGCGTTGCCACACCCCATCTCCC 180

QY 215 GTACCAACATCTCCCTCTGTAAACCCCTCGGCTGGGCTCTCCCTCCCT 274

DB 181 GTACCAACATCTCCCTCTGTAAACCCCTCGGCTGGGCTCTCCCTCCCT 240

QY 275 CTCCATATCAGTGAATCTTCCAGTCTTTCGAGGCTCTGGGCTCTCCCTCCCT 334

DB 241 CTCCATATCAGTGAATCTTTCGAGTCTTTCGAGGCTCTGGGCTCTCCCTCCCT 300

QY 335 GGGCTTTCACACCTCCCTCTATGGGCTCTATGAGGCTCTGAGGCTCTGAGATTAA 394

DB 301 GGGCTTTCACACCTCCCTCTATGGGCTCTATGAGGCTCTGAGGCTCTGAGATTAA 360

QY 395 AACCTGGTTCGATGCTGAATGAAGACCGTAAGACCAAGCAAGCAAGCACTGT 454

DB 361 AACCTGGTTCGATGCTGAATGAAGACCGTAAGACCAAGCAAGCAAGCACTGT 420

QY 455 CTCTGCTGCTGATATCCTTCACACCTGGGAACATCCCAAGACACCTCTTAACTCG 514

DB 421 CTCTGCTGCTGATATCCTTCACACCTGGGAACATCCCAAGACACCTCTTAACTCG 480

QY 515 GAGCAGAGATGCTGGGAGAGCTGGGGCGGCTGGCTGTACTTGAAGTCTGAGA 574

DB 481 GAGCAGAGATGCTGGGAGAGCTGGGGCGGCTGGCTGTACTTGAAGTCTGAGA 540

QY 575 AGAGAGAGCTGAAGAGAGTTCAGTCTTCGCAATTAAGCGCACTCAGAGAGCTTT 634

DB 541 AGAGAGAGCTGAAGAGAGTTCAGTCTTCGCAATTAAGCGCACTCAGAGAGCTTT 600

QY 635 CGGCTGAGACACCCCGTCAAGCCAGAGAAACAGATGAGTGGGCTCTGACTGG 694

DB 601 CGGCTGAGACACCCCGTCAAGCCAGAGAAACAGATGAGTGGGCTCTGACTGG 660

QY 695 TGGCTCAGATGAGGAGAGCGGAGCTGGGAGCTTACCTCTCACTAGGAGAGATGG 754

DB 661 TGGCTCAGATGAGGAGAGCGGAGCTGGGAGCTTACCTCTCACTAGGAGAGATGG 720

QY 755 GGGTGAAGTCACTGTGGGCCCAAGGCCAGAAAGGGGAGGCCACTCTCCCTATTCCT 814

DB 721 GGGTGAAGTCACTGTGGGCCCAAGGCCAGAAAGGGGAGGCCACTCTCCCTATTCCT 780

QY 815 ACAGCCCAAGTGAACCCCACTGGGGTCTCCCAAGCCACCCACTCCAGAGTCTAA 874

DB 781 ACAGCCCAAGTGAACCCCACTGGGGTCTCCCAAGCCACCCACTCCAGAGTCTAA 840

QY 875 TGGCTGATTCATGATTTGGCGAGGGGTGACCCAGGAGCTCAGAGAGAGGTTTGA 934

DB 841 TGGCTGATTCATGATTTGGCGAGGGGTGACCCAGGAGCTCAGAGAGAGGTTTGA 900

QY 935 GACAGCTGCTGACATCTGGAACCGCGCTGAGAGAAATCTTGTCTCACTCTCTAC 994

DB 901 GACAGCTGCTGACATCTGGAACCGCGCTGAGAGAAATCTTGTCTCACTCTCTAC 960

QY 995 AAGCTTTCAGAGCTCCCAAGACATGAGTCTTCAAGCCAGAGTCAACCAAGCCCA 1054

DB 961 AAGCTTTCAGAGCTCCCAAGACATGAGTCTTCAAGCCAGAGTCAACCAAGCCCA 1020

QY 1055 CATCCACAGAGTCTGGGAGAGTGGAGATCCCACTCAGCCAGCTTACAGCCAGAG 1114

DB 1021 CATCCACAGAGTCTGGGAGAGTGGAGATCCCACTCAGCCAGCTTACAGCCAGAG 1080

QY 1115 AGCAGAGGCTCTTGGAGCCCAATGGCTCTGATGAAAGTCAAGAAATTTACTACAG 1174

DB 1081 AGCAGAGGCTCTTGGAGCCCAATGGCTCTGATGAAAGTCAAGAAATTTACTACAG 1140

QY 1175 AATCAGAAAGAGAGAGAGAAATCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1234

DB 1141 AATCAGAAAGAGAGAGAGAAATCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1200

QY 1235 TAGGAAGCCGCCCAAGGAGACACAGGCTCAGGCCACCAAGAGAGAGAGAGAG 1294

DB 1201 TAGGAAGCCGCCCAAGGAGACACAGGCTCAGGCCACCAAGAGAGAGAGAGAG 1260

QY 1295 CTGTGAGAGAGAGCTCTGTTCACATGAGCCCTGAGAAATGAGATTTTAACTAA 1354

DB 1261 CTGTGAGAGAGAGCTCTGTTCACATGAGCCCTGAGAAATGAGATTTTAACTAA 1320

QY 1355 TCACAGAGTCTACTTTCACAAAGAGCTCACCCAGAGAGAGAGAGAGAGAGAG 1414

DB 1321 TCACAGAGTCTACTTTCACAAAGAGCTCACCCAGAGAGAGAGAGAGAGAGAG 1380

QY 1415 GAAGCTGAGCTGATTTATGAGAGAGAGATCAGAGATTTATGAGATCAGAGCTAT 1474

DB 1381 GAAGCTGAGCTGATTTATGAGAGAGAGATCAGAGATTTATGAGATCAGAGCTAT 1440

QY 1475 TTGGCCAGGCTGATATCCCAAGAACTTCGATATGATCTCAGGGGCTGTGAA 1534

DB 1441 TTGGCCAGGCTGATATCCCAAGAACTTCGATATGATCTCAGGGGCTGTGAA 1500

QY 1535 TTGGGAAGTCAACATCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594

DB 1501 TTGGGAAGTCAACATCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560

QY 1595 GGGAGCGCTTCCAGATGCTTCTTACTTCACTGAGTCAAGAGCTGAGCTCAAGAGTGG 1654

DB 1561 GGGAGCGCTTCCAGATGCTTCTTACTTCACTGAGTCAAGAGCTGAGCTCAAGAGTGG 1620

QY 1655 TGAATCTCGTGAAGCTATCGAGAAAGATGAGAGACAGCACTCGGCTCCCATTAAGACA 1714

DB 1621 TGAATCTCGTGAAGCTATCGAGAAAGATGAGAGACAGCACTCGGCTCCCATTAAGACA 1680

QY 1715 TCCTGTCTAGGAG 1774


```
Db 1681 TCCCTGCTAGGCCAGAGCGGCTGCTCTTCATCTCCATGCTGTATGATGAGCCAGATGGG 1740
Qy 1775 TCTTTCAGAGGCGGAGTTCTGAGCTGTGTCTGCACTGGAGCCAGCCACAGCGGCGGATG 1834
Db 1741 TCTTTCAGAGGCGGAGTTCTGAGCTGTGTCTGCACTGGAGCCAGCCACAGCGGCGGATG 1800
Qy 1835 CATGCTGGGCGAGTTTGTCTGGGGAAAATCTATCTCCCGAGGCACTCCTTCTGATCAAGG 1894
Db 1801 CACTGCTGGGCGAGTTTGTCTGGGGAAAATCTATCTCCCGAGGCACTCCTTCTGATCAAGG 1860
Qy 1895 CTGGAGCAAGCTCTGCAAACTCATCTCTTCTTGGAGCAGGCACTTGGGTAGAGG 1954
Db 1861 CTGGAGCAAGCTCTGCAAACTCATCTCTTCTTGGAGCAGGCACTTGGGTAGAGG 1920
Qy 1955 TCTTGGGGTCTCTGATGTCAGAGAGAAATATTTCTACAGATATTTTCAAGATGAAA 2014
Db 1921 TCTTGGGGTCTCTGATGTCAGAGAGAAATATTTCTACAGATATTTTCAAGATGAAA 1980
Qy 2015 GGGAGCAATTAAGCTTTAGGTTGGTCAATCAAAAGAGCTTGGGCTGTGTC 2074
Db 1981 GGGAGCAATTAAGCTTTAGGTTGGTCAATCAAAAGAGCTTGGGCTGTGTC 2040
Qy 2075 TTGTGCTGGGTGTCTGCTGGCTGGCTGCACTTGGCTGATGACAGATGAAGCGAGG 2134
Db 2041 TTGTGCTGGGTGTCTGCTGGCTGGCTGCACTTGGCTGATGACAGATGAAGCGAGG 2100
Qy 2135 AAAAATCACACTGACTTCCAGAACCAACAACCTCTGTCTATCAATTACCTTGCAGG 2194
Db 2101 AAAAATCACACTGACTTCCAGAACCAACAACCTCTGTCTATCAATTACCTTGCAGG 2160
Qy 2195 CTCTCCAGGTCAGCACTTGGAGCCCAAGTCCAGAGACCTCTGCTCTGGCTGTAGG 2254
Db 2161 CTCTCCAGGTCAGCACTTGGAGCCCAAGTCCAGAGACCTCTGCTCTGGCTGTAGG 2220
Qy 2255 GCATCTGGCAAAAAGAACCTTTTCTGAGTCAGATGACTCAAGAAAGCATGGGTAGATG 2314
Db 2221 GCATCTGGCAAAAAGAACCTTTTCTGAGTCAGATGACTCAAGAAAGCATGGGTAGATG 2280
Qy 2315 GGGCCATCATCTCCACCTTTTGAAGTGGTATTTCTTCAAGAGCACCCCATCTCTGA 2374
Db 2281 GGGCCATCATCTCCACCTTTTGAAGTGGTATTTCTTCAAGAGCACCCCATCTCTGA 2340
Qy 2375 GCTACAGCTTCAATTCACCTGTGTTCCAAAGTCTTGGAGCAATCTCTATGTCTTGG 2434
Db 2341 GCTACAGCTTCAATTCACCTGTGTTCCAAAGTCTTGGAGCAATCTCTATGTCTTGG 2400
Qy 2435 AGGATGAGAGGGGAGAGTAAACATTTCTAATTGCATATAGATTGGAAAAGCGTAG 2494
Db 2401 AGGATGAGAGGGGAGAGTAAACATTTCTAATTGCATATAGATTGGAAAAGCGTAG 2460
Qy 2495 AAGCATATGGAATATCATGCGCTGTGTTGGGGCATCAACACAGCTTCTTATGGGCTGT 2554
Db 2461 AAGCATATGGAATATCATGCGCTGTGTTGGGGCATCAACACAGCTTCTTATGGGCTGT 2520
Qy 2555 TAAGTATAGGGGGAGAGAGATGGAACAATCTTTCATCTGCGGCTGTCTCAGGGGA 2614
Db 2521 TAAGTATAGGGGGAGAGAGATGGAACAATCTTTCATCTGCGGCTGTCTCAGGGGA 2580
Qy 2615 GGAACCTGATGAGTGGGTCCCGTCCCTGAGCTGTGTCAGGCACTCTCTGGAGT 2674
Db 2581 GGAACCTGATGAGTGGGTCCCGTCCCTGAGCTGTGTCAGGCACTCTCTGGAGT 2640
Qy 2675 CCGTCACTGCTTGTACGAGACTTGGAACAAAAGTTCTGTACACAAGTATGGGCCATT 2734
Db 2641 CCGTCACTGCTTGTGTAGACTCGGAACAAGTTCTGTACACAAGTATGGGCCATT 2700
Qy 2735 TCGAAGAAATGGGAGTGTGTAGAAACAGACATGAGCTCTTATGTGTGACTTTCTGCA 2794
Db 2701 TCGAAGAAATGGGAGTGTGTAGAAACAGACATGAGCTCTTATGTGTGACTTTCTGCA 2760
Qy 2795 TTAATTCAGCGCCAGCGTGAAGAGCTTCACTGATTTGAGGGCAGGCAAGATCAA 2854
Db 2761 TTAATTCAGCGCCAGCGTGAAGAGCTTCACTGATTTGAGGGCAGGCAAGATCAA 2820
Qy 2855 CATGAGCCCAACCATGATGATCTCTGTTCAAGTGGGTCCTCCAGTCAAGATGCTTATGGC 2914
Db 2821 CATGAGCCCAACCATGATGATCTCTGTTCAAGTGGGTCCTCCAGTCAAGATGCTTATGGC 2880
Qy 2915 AGATTCCTTCTCCGCTCCCAAGGTCACAGAAACCTGAAAGAGCTGGACCTTAAGTGA 2974
Db 2881 AGATTCCTTCTCCGCTCCCAAGGTCACAGAAACCTGAAAGAGCTGGACCTTAAGTGA 2940
Qy 2975 ACTGCTGAGCCACTCTGCAAGTGAAGAGTCTTTGTAAAGACCTGAGACGCTCTGCTGC 3034
Db 2941 ACTGCTGAGCCACTCTGCAAGTGAAGAGTCTTTGTAAAGACCTGAGACGCTCTGCTGC 3000
Qy 3035 TCTTGAAGACCTTGGGTGTGCTGTGCTGTGCTCAACAGTGAAGACTGCAAGACCTTG 3094
Db 3001 TCTTGAAGACCTTGGGTGTGCTGTGCTGTGCTCAACAGTGAAGACTGCAAGACCTTG 3060
Qy 3095 CCGTTGGGCTGAGAGCAACAGACCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCA 3154
Db 3061 CCGTTGGGCTGAGAGCAACAGACCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCA 3120
Qy 3155 CGGATGCTGAGAGCAACACTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAG 3214
Db 3121 CGGATGCTGAGAGCAACACTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAG 3180
Qy 3215 GACTGAGCTGTGCTGCTGTGCTGCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3274
Db 3181 GACTGAGCTGTGCTGCTGTGCTGCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 3275 TTATGTCAGGCCCCAGCTGAAGAGCTGACCTGAGAGCAACCTGAGATGAGTGG 3334
Db 3241 TTATGTCAGGCCCCAGCTGAAGAGCTGACCTGAGAGCAACCTGAGATGAGTGG 3300
Qy 3335 GCGTGGCACTGCTCTGTGAGGGGCTCAGGCACTCTGCTCAAACTCATAGCTGGGCTG 3394
Db 3301 GCGTGGCACTGCTCTGTGAGGGGCTCAGGCACTCTGCTGCAAACTCATAGCTGGGCTG 3360
Qy 3395 TGGACCAACAACCTCTGAGTATGATGATGAGAGCAAGACTGAGGGCTCTGGAGCAGAGA 3454
Db 3361 TGGACCAACAACCTCTGAGTATGATGATGAGAGCAAGACTGAGGGCTCTGGAGCAGAGA 3420
Qy 3455 AACCTGAGCTGCTCATCTTCAAGAGCGGAAACCAAGTATGAGACCCCTCATGAGGGCT 3514
Db 3421 AACCTGAGCTGCTCATCTTCAAGAGCGGAAACCAAGTATGAGACCCCTCATGAGGGCT 3480
Qy 3515 TGGATACGGGAGAGATGATTAATAGCAATCTCACTCAAGCGGAGAGACTGAGTACG 3574
Db 3481 TGGATACGGGAGAGATGATTAATAGCAATCTCACTCAAGCGGAGAGACTGAGTACG 3540
Qy 3575 AGAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACCTCTGAGACGTGACAGATCT 3634
Db 3541 AGAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACCTCTGAGACGTGACAGATCT 3600
Qy 3635 TCCCAATGCTGAGATTGAGAGGAAAGCTCCCAAGGATGATGACCGGTGGAATCTGTGT 3694
Db 3601 TCCCAATGCTGAGATTGAGAGGAAAGCTCCCAAGGATGATGACCGGTGGAATCTGTGT 3660
Qy 3695 GCGTGCCTTCTCTGCTCTCAAGGGAGCTGTGATAGGAAGCTTTTGGGAGCTGACGATG 3754
Db 3661 GCGTGCCTTCTCTGCTCTCAAGGGAGCTGTGATAGGAAGCTTTTGGGAGCTGACGATG 3720
Qy 3755 ACTTCTGGGGCCCAAGGGGCTGTGGCTTATCTGAGGATGATGCAAAAGAAAGAACTTGT 3814
Db 3721 ACTTCTGGGGCCCAAGGGGCTGTGGCTTATCTGAGGATGATGCAAAAGAAAGAACTTGT 3780
Qy 3815 ACCGAGTTCATCTCCCTGTAGCTGCTCAACCGCTGAGCCCAACAGGGGCTCTGCTTGG 3874
Db 3781 ACCGAGTTCATCTCCCTGTAGCTGCTCAACCGCTGAGCCCAACAGGGGCTCTGCTTGG 3840
Qy 3875 TGATGAGAGAGCGGTGACCGTTGATGATTAATCTGTGTGTGGGACAGTTCTGGGTG 3934
Db 3841 TGATGAGAGAGCGGTGACCGTTGATGATTAATCTGTGTGTGGGACAGTTCTGGGTG 3900
```


QY	3335	AGATCAACCCACACACACAGCTGGAATGGGTGGCAAGGACCTGTGTGACATCAAGGCTAGC	3394
Db	3301	AGATCAACCCACACACACAGCTGGAATGGGTGGCAAGGACCTGTGTGACATCAAGGCTAGC	3360
QY	3395	CTGAGAGCTGTGGAAGCTGTGACCTCCCTCACTTTGTGTGCTCTCCAAAGGGGGCATGTGG	4054
Db	3361	CTGAGAGCTGTGGAAGCTGTGACCTCCCTCACTTTGTGTGCTCTCCAAAGGGGGCATGTGG	4020
QY	4055	ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAAAGCCAG	4114
Db	4021	ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAAAGCCAG	4080
QY	4115	CCAGGGTGGAGCTCATCACTATGTTCTGAAAAACCCAGGTTCTCCGCCCTTGGAGATGC	4174
Db	4081	CCAGGGTGGAGCTCATCACTATGTTCTGAAAAACCCAGGTTCTCCGCCCTTGGAGATGC	4140
QY	4175	TCTGAAAATGATTCATATATGCCCTGTGGGCTTCATTCCCTGACCTGTGTGTGCTTT	4234
Db	4141	TCTGAAAATGATTCATATATGCCCTGTGGGCTTCATTCCCTGACCTGTGTGTGCTTT	4200
QY	4235	ACCAACGGGTCATCTCTGAGGAATGCATCCATCCACTACCTGATCCGAATCCAAATGACGTCT	4294
Db	4201	ACCAACGGGTCATCTCTGAGGAATGCATCCATCCACTACCTGATCCGAATGACGTCT	4260
QY	4295	CCATTTCGG-----	4302
Db	4261	CCATTTCGGAGGCGCATATGATCTAGAAATGAATAATCAGTTTGTGGATTCACAAGC	4320
QY	4303	-----	4302
Db	4321	CACCCCGCTGACCCCACTTATATATGAGGCTGTGCTTACACTGTGTCTGGTCTGTTCAG	4380
QY	4303	-----AAGAACTGAGACTCTGCTATCAAGCCCTTGGAGAAAGCC	4342
Db	4381	GGATGCTGGAATATCTCCCAAGGAATGGAAGCTCTGTATCGAAAGCCTTGGAGAAAGCC	4440
QY	4343	AGCTGTTCTGGAATTCTACGTTTGCCCATCTTGGATCAGGGATCAGGCTGCAATGAAAG	4402
Db	4441	AGCTGTTCTGGAATCTCTACGTTTGCCCATCTTGGATCAGGGATCAGGCTGCAATGAAAG	4500
QY	4403	ACAGAAGAGATGACATCTGGTGTGGGAGGCTTGTGTGAACACAGAGAGATCTATGTCCTG	4462
Db	4501	ACAGAAGAGATGACATCTGGTGTGGGAGGCTTGTGTGAACACAGAGAGATCTATGTCCTG	4560
QY	4463	CAACTACTGTATCCCTCCAGCCCGCATAGCCGATACCTTCACTCTGTGATGCCCGCAGT	4522
Db	4561	CAACTACTGTATCCCTCCAGCCCGCATAGCCGATACCTTCACTCTGTGATGCCCGCAGT	4620
QY	4523	TGCTGCACCTTTGTGAGCAATATGAGAGACGCTGATATGCCGATGACATCGGTGAGG	4582
Db	4621	TGCTGCACCTTTGTGAGCAATATGAGAGACGCTGATATGCCGATGACATCGGTGAGG	4680
QY	4583	TTGTCTTGGACAACACTGATGAGCAGGAGCTGAGACCGAGACAGATACGAGAGGAGTGTGCG	4642
Db	4681	TTGTCTTGGACAACACTGATGAGCAGGAGCTGAGACCGAGACAGATACGAGAGGAGTGTGCG	4740
QY	4643	CTGAGAACACGAGGCTCAGCCAGATGCGGAAGCTGTTCACTTGAAGCCAGTCTTGGAGCC	4702
Db	4741	CTGAGAACACACGAGGCCACAGATGCGGAAGCTGTTCACTTGAAGCCAGTCTTGGAGCC	4800
QY	4703	GGAGGTGCAAAAGATGAGACTTACCAAGGCCCTGAAAGGAGAACCAATCTCACTCATTTATGG	4762
Db	4801	GGAGGTGCAAAAGATGAGACTTACCAAGGCCCTGAAAGGAGAACCAATCTCACTCATTTATGG	4860
QY	4763	AACTCTGGAGAAAGGCGACAAAAAAGGAGCTCTGCGCATCTGACAGAGCTGAAGTATCAACA	4822
Db	4861	AACTCTGGAGAAAGGCGCGACAAAAAAGGAGCTCTGCGCATCTGACAGAGCTGAAGTATCAACA	4920
QY	4823	CCAGCCCTTGAACCTTTAGTCTCTGGCTTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTT	4882
Db	4921	CCAGCCCTTGAACCTTTAGTCTCTGGCTTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTT	4980

Qy		4983	CTCTGCAAAACAAGTGGCATCTGTGGTTGCCTTCACAGACTAAAGTAATGGAACCTTGATG	4942
Db		4981	CTCTGCAAAACAAGTGGCATCTGTGGTTGCCITTCACAGACTAAAGTAATGGAACCTTAGAT	5040
Qy		4943	ATGCCCTTGTGGGCATTATGTGTCCATGCCAGGGAATGCCACAGGGGGCC	4992
Db		5041	GATGCCCTTGTGGGCATTATGTGTCCA-GCCAGGGGAATGCCACAGGGGGCCC	5089
<hr/>				
RESULT 6				
AAD02760	ID	AAD02760	standard; cDNA; 4422 BP.	
XX	AC			
XX	ADD02760;			
DT				
DS	31-MAY-2001	(first entry)		
XX				
DS	Human NB-ARC and CARD containing protein (NAC) beta isoform cDNA.			
XX				
KW	Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;			
KM	caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;			
KW	cysteine aspartyl protease; apoptosis; cytokine production;			
KM	cytokine receptor signalling; therapy; inflammatory disorder; sepsis;			
KX	fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	1..4422		
FT		/tag= a		
FT		/product= "Human NB-ARC and CARD containing protein (NAC)		
FT		beta isoform"		
FT		/transl_except= (pos:3892..3915, aa:RDAGNTDQ)		
FT		/note= "The translation exceptions only occur when		
FT		decoding the alternate version of the protein (AAY72711)"		
FT		2869..2961		
FT	misc_feature	/tag= b		
FT		/note= "this region is not found in human NAC gamma		
FT		isoform (AAD02761) and human NAC delta isoform		
FT		(AAD02762) due to the alternative mRNA splicing"		
FT		3781..3915		
FT	misc_feature	/tag= c		
FT		/note= "this region is not found in human NAC gamma		
FT		isoform (AAD02761) due to the alternative mRNA splicing"		
FT				
PN	MO200116170-A2.			
PD	08-MAR-2001.			
PF	01-SEP-2000; 2000MO-US24152.			
PR	01-SEP-1999; 99US-0388221.			
PA	(BURN-) BURNHAM INST.			
PI	Reed JC;			
DR	WP1; 2001-183258/19.			
DR	P-PADB; AAY72669, AAY72711.			
PT	Novel nucleic acid encoding NB-ARC and caspase associated recruitment			
PT	domains, used to produce polypeptides for screening for modulators of			
PT	apoptosis -			
PS	Claim 4; Page 127-133; 184pp; English.			
CC	The present sequence is a human NB-ARC and CARD containing protein			
CC	(NAC) beta isoform cDNA. NAC beta isoform represents the NAC splice			
CC	variant in which both the splice regions are present in the translated			
CC	polypeptide. NAC protein comprises a nucleotide binding (NB) domain			
CC	(also referred as NB-ARC domain), a caspase-associated recruitment			
CC	domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine			
CC	aspartyl proteases, are principal effectors of apoptosis. CARD containing			

CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.

XX Sequence 4422 BP; 1043 A; 1224 C; 1240 G; 915 T; 0 other;

Query Match 76.2%; Score 4148; DB 22; Length 4422;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

QY 523 ATGGCTGGCCGAGGCTTGGGCGCCGCTGGCTGTTACTTGGAGTTCTGGAAGAGAGAGAG 582
 DB 1 ATGGCTGGCCGAGGCTTGGGCGCCGCTGGCTGTTACTTGGAGTTCTGGAAGAGAGAGAG 60
 QY 583 CTGAAGAGAGTTCCAGCTTCTGCTGCGCAATTAAGCGCACTCCAGAGAGCTCTGGGGTGA 642
 DB 61 CTGAAGAGAGTTCCAGCTTCTGCTGCGCAATTAAGCGCACTCCAGAGAGCTCTGGGGTGA 120
 QY 643 ACAACCGCTCAGCCAG 702
 DB 121 ACAACCGCTCAGCCAG 180
 QY 703 TATGGGAG 762
 DB 181 TATGGGAG 240
 QY 763 TCACTGTCGCGCCAG 822
 DB 241 TCACTGTCGCGCCAG 300
 QY 823 AGTGAACCCCACTGGGGGTCTCCAGCCAACTCCAGCCAGAGAGAGAGAGAGAGAGAG 882
 DB 301 AGTGAACCCCACTGGGGGTCTCCAGCCAACTCCAGCCAGAGAGAGAGAGAGAGAGAG 360
 QY 883 ATCCATGAAATTTGGCGGGGGGTGCAACCAAGGGCTCAAGAGAGAGAGAGAGAGAGAG 942
 DB 361 ATCCATGAAATTTGGCGGGGGGTGCAACCAAGGGCTCAAGAGAGAGAGAGAGAGAGAG 420
 QY 943 CTTGACACATCTGGAACGCGCTGGAGAGAAATCTCTGCTCACTCTCTCAAGAGCTCT 1002
 DB 421 CTTGACACATCTGGAACGCGCTGGAGAGAAATCTCTGCTCACTCTCTCAAGAGCTCT 480
 QY 1003 CCAAGCTCCCAAGCACTAGAGTCTCAAGCCAGAGAGTCAACCAAGCCCAATCCACA 1062
 DB 481 CCAAGCTCCCAAGCACTAGAGTCTCAAGCCAGAGAGTCAACCAAGCCCAATCCACA 540
 QY 1063 GCAAGTCTGGGAGAGCTGGGGATCCCACTCAAGCCAGCTTACCAACCAAGAGAGAGAG 1122
 DB 541 GCAAGTCTGGGAGAGCTGGGGATCCCACTCAAGCCAGCTTACCAACCAAGAGAGAGAG 600
 QY 1123 GCTCTCTGGAGCCCAATGAGCTTGGAGAAACGTCAAGAAATTTACTACAGAGAAATCAGA 1182
 DB 601 GCTCTCTGGAGCCCAATGAGCTTGGAGAAACGTCAAGAAATTTACTACAGAGAAATCAGA 660
 QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCATGGGAGAGAGAGAGAGAG 1242
 DB 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCATGGGAGAGAGAGAGAGAG 720
 QY 1243 CCCCCACAGGCGCACACAGCTTACAGCCCAACCAACCAAGAGAGAGAGAGAGAGAGAG 1302
 DB 721 CCCCCACAGGCGCACACAGCTTACAGCCCAACCAACCAAGAGAGAGAGAGAGAGAGAG 780
 QY 1303 GAG 1362
 DB 781 GAG 840
 QY 1363 CTGCTACTTCAACAG 1422
 DB 841 CTGCTACTTCAACAG 900

QY 1423 CTTGATTTGTGAG 1482
 DB 901 CTTGATTTGTGAG 960
 QY 1483 GGCCTGATATCCCAAG 1542
 DB 961 GGCCTGATATCCCAAG 1020
 QY 1543 TCAACCTGGCCAG 1602
 DB 1021 TCAACCTGGCCAG 1080
 QY 1603 TTCCAGAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1662
 DB 1081 TTCCAGAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1140
 QY 1663 GCTGAGTCACTGGAG 1722
 DB 1141 GCTGAGTCACTGGAG 1200
 QY 1723 AGGCCAG 1782
 DB 1201 AGGCCAG 1260
 QY 1783 GAGCCAG 1842
 DB 1261 GAGCCAG 1320
 QY 1843 GGCAGTTTGTGGAG 1902
 DB 1321 GGCAGTTTGTGGAG 1380
 QY 1903 ACAGCTCTGAG 1962
 DB 1381 ACAGCTCTGAG 1440
 QY 1963 TTCTGAG 2022
 DB 1441 TTCTGAG 1500
 QY 2023 ATTAG 2082
 DB 1501 ATTAG 1560
 QY 2083 TGGGTGTCTGGAG 2142
 DB 1561 TGGGTGTCTGGAG 1620
 QY 2143 ACACTGATCTCCAG 2202
 DB 1621 ACACTGATCTCCAG 1680
 QY 2203 GCTCAG 2262
 DB 1681 GCTCAG 1740
 QY 2263 GAAAAAG 2322
 DB 1741 GAAAAAG 1800
 QY 2323 ATCTCAG 2382
 DB 1801 ATCTCAG 1860
 QY 2383 TTCAATCACTCTGTTTCAAG 2442
 DB 1861 TTCAATCACTCTGTTTCAAG 1920
 QY 2443 AAGGAG 2502
 DB 1921 AAGGAG 1980
 QY 2503 GGAATACATGAG 2562

```
1981 GGAATACATGGCTGTTTGGGGCATCAACACGTTTCTATTGGGCCCTGTTAAATGAT 2040
QY 2563 GAGGGGAGAGAGAGATGAGAAATCTTTCTACCTGCCGCTGTCTCAGGGGAGAACTTG 2622
Db 2041 GAGGGGAGAGAGAGATGAGAAATCTTTCTACCTGCCGCTGTCTCAGGGGAGAACTTG 2100
QY 2623 ATGCAGTGGGTCCCGTCCCTGAGCTGCTGCTGACGACACCTCTGAGTCCCTCCAC 2682
Db 2101 ATGCAGTGGGTCCCGTCCCTGAGCTGCTGCTGACGACACCTCTGAGTCCCTCCAC 2160
QY 2683 TGCCTTACAGACTCGGAACTGAACTTCTGACACAACTGATGGCCCTTTGAGAA 2742
Db 2161 TGCCTTACAGACTCGGAACTGAACTTCTGACACAACTGATGGCCCTTTGAGAA 2220
QY 2743 ATGGGATGTGTGTAGAAACAGACATGAGACTCTTATGTGTGACTTTTGCATTAATTC 2802
Db 2221 ATGGGATGTGTGTAGAAACAGACATGAGACTCTTATGTGTGACTTTTGCATTAATTC 2280
QY 2803 AGCCGCGAGTGAAGAGCTTCACTGATTGAGGGGAGGAGACAGATCAATGAGAGC 2862
Db 2281 AGCCGCGAGTGAAGAGCTTCACTGATTGAGGGGAGGAGACAGATCAATGAGAGC 2340
QY 2863 CCCACCATGTAGTCTCTGTTCAAGTGGGTCCAGTCAAGATGCTATTGGCAGATTCTC 2922
Db 2341 CCCACCATGTAGTCTCTGTTCAAGTGGGTCCAGTCAAGATGCTATTGGCAGATTCTC 2400
QY 2923 TTCTCCGTCTCAAGTCAACAGAACTTGAAGAGCTGAGCTTAAATGTGAACTCGCTG 2982
Db 2401 TTCTCCGTCTCAAGTCAACAGAACTTGAAGAGCTGAGCTTAAATGTGAACTCGCTG 2460
QY 2983 AGCCACTCTGCAATGAAGAGCTTTGTAAAGACCTGAGAGCCCTGCTGCTCCCTGAGG 3042
Db 2461 AGCCACTCTGCAATGAAGAGCTTTGTAAAGACCTGAGAGCCCTGCTGCTCCCTGAGG 2520
QY 3043 ACCCTGGGTGTGGCTGTGCTGTGCTCAAGTGAAGACTGCAAGGACCTTGGCTTGGG 3102
Db 2521 ACCCTGGGTGTGGCTGTGCTGTGCTCAAGTGAAGACTGCAAGGACCTTGGCTTGGG 2580
QY 3103 CTGAGAGCCAAACAGACCTGACCTGAGCTGAGCTTGAATGCTCTCAAGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTGACCTGAGCTGAGCTTGAATGCTCTCAAGATGCT 2640
QY 3163 GAGAGCCAAACACCTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTCAAGCACTGCG 3222
Db 2641 GAGAGCCAAACACCTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTCAAGCACTGCG 2700
QY 3223 CTGGTCAAGTGTGGCTCAAGTCTGACTGTGCTGCTGCAAGGACTGGCTCTGTGCTTAA 3282
Db 2701 CTGGTCAAGTGTGGCTCAAGTCTGACTGTGCTGCTGCAAGGACTGGCTCTGTGCTTAA 2760
QY 3283 AGCCCAAGCTTGAAGAGCTTGAACCTGACAGAAACAACCTGATGACGTTGGCGTGGCA 3342
Db 2761 AGCCCAAGCTTGAAGAGCTTGAACCTGACAGAAACAACCTGATGACGTTGGCGTGGCA 2820
QY 3343 CTGCTCTGTAGAGGGGCTCAGGCACTCCGCTGCAAAATCTAATAGCCCTGGGGCTGAG 3402
Db 2821 CTGCTCTGTAGAGGGGCTCAGGCACTCCGCTGCAAAATCTAATAGCCCTGGGGCTGAG 2880
QY 3403 ACAACTCTGATGATGATGAGGAGGAACTGAGGGCCCTGAGAGCAGAGAAACCTCAG 3462
Db 2881 ACAACTCTGATGATGATGAGGAGGAACTGAGGGCCCTGAGAGCAGAGAAACCTCAG 2940
QY 3463 CTGCTCATCTTCAAGCAGACGAGAAACAAGTGTGATGACCCCTTACTGAGGGCTTGA 3522
Db 2941 CTGCTCATCTTCAAGCAGACGAGAAACAAGTGTGATGACCCCTTACTGAGGGCTTGA 3000
QY 3523 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3582
Db 3001 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
QY 3583 GCTTCCCATGTGTGCTCAAGGCTAATCTCAACTCTGAGCTGAGCAAGATCTTCCCAATT 3642
3583 GCTTCCCATGTGTGCTCAAGGCTAATCTCAACTCTGAGCTGAGCAAGATCTTCCCAATT 3642

3061 GCTTCCCATGTGTGCTCAAGGCTAATCTCAAACTCTGAGAGCTGAGCAAGATCTTCCCAATT 3120
QY 3643 GCTGAGATTTGCAAGAGAAAGCTCCCAAGAGTGAATCCGCTGAAACTCTTGTGCGTCC 3702
Db 3121 GCTGAGATTTGCAAGAGAAAGCTCCCAAGAGTGAATCCGCTGAAACTCTTGTGCGTCC 3180
QY 3703 TCTCCGCTCTCAAGGGGAGCTGCAATCGAAGCCCTTGGGGGAGCTGAGATGACTCTG 3762
Db 3181 TCTCCGCTCTCAAGGGGAGCTGCAATCGAAGCCCTTGGGGGAGCTGAGATGACTCTG 3240
QY 3763 GGGCCCAAGGGGCTGTGCTGCTAAGAGTGTGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3241 GGGCCCAAGGGGCTGTGCTGCTAAGAGTGTGACAAAGAAAGAACTTGTACCGAGTT 3300
QY 3823 CACTTCCCTGTAGCTGTGCTCTCAAGCTGAGCCCAACAGGGGTCTGCTTGTGTATGAA 3882
Db 3301 CACTTCCCTGTAGCTGTGCTCTCAAGCTGAGCCCAACAGGGGTCTGCTTGTGTATGAA 3360
QY 3883 GAAAGGATGACCGTTGAGATTGAATTTGTGTGTGTGAGGACAGTTCCTGGGTGAGTCAAC 3942
Db 3361 GAAAGGATGACCGTTGAGATTGAATTTGTGTGTGTGAGGACAGTTCCTGGGTGAGTCAAC 3420
QY 3943 CCACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4002
Db 3421 CCACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3480
QY 4003 GTGAGAGCTGTGACACTCTCTCACTTTGTGCTCTTCAAGGGGGCCATGTGACATCC 4062
Db 3481 GTGAGAGCTGTGACACTCTCTCACTTTGTGCTCTTCAAGGGGGCCATGTGACATCC 3540
QY 4063 CTGTTCCAAATGAGCCCACTTTAAAGAGAGGGAGTGTCTCTGAGAGAGCAGCAGAGGTG 4122
Db 3541 CTGTTCCAAATGAGCCCACTTTAAAGAGAGGGAGTGTCTCTGAGAGAGCAGCAGAGGTG 3600
QY 4123 GAGCTGACATCAATGATTTGAGAAACCCCAAGCTTCTCCCTTGGGAGTCTCTGAAA 4182
Db 3601 GAGCTGACATCAATGATTTGAGAAACCCCAAGCTTCTCCCTTGGGAGTCTCTGAAA 3660
QY 4183 ATGATTCATTAATGCTCTGCTGCTTCAATCCCGTCACTCTGTGTGTGTGCTTTAAAC 4242
Db 3661 ATGATTCATTAATGCTCTGCTGCTTCAATCCCGTCACTCTGTGTGTGTGCTTTAAAC 3720
QY 4243 GTTCATCTGAGAGAGTCACTTCCACTCTCACTGATCCCAAGTGAAGCTGCTCATTTCCG 4302
Db 3721 GTTCATCTGAGAGAGTCACTTCCACTCTCACTGATCCCAAGTGAAGCTGCTCATTTCCG 3780
QY 4303 ----- 4302
Db 3781 AAGGCCATGATGATCTAGAAATGAATTCAGTTTGTGCAATCCAAAGCCACCCCG 3840
QY 4303 ----- 4302
Db 3841 CTGACCCCACTTTAATATGAGCTGTGCTTAACTGTGTGTGAGCTTCAAGGATGCTG 3900
QY 4303 ----- AAGGAACTGAGAGCTCTGCTTCAAGAGCCCTGAGAGAACCAAGCTGCTC 4350
Db 3901 GAAATATCTCCCAAGGAACTGAGAGCTCTGCTTCAAGAGCCCTGAGAGAACCAAGCTGCTC 3960
QY 4351 TCGAGATTCTAGTTGGCCACTTGGGATCAGGAGATCAGGCTCAAGTGAAGCAAGAAA 4410
Db 3961 TCGAGATTCTAGTTGGCCACTTGGGATCAGGAGATCAGGCTCAAGTGAAGCAAGAAA 4020
QY 4411 GATGAGACTGTGTGTGTGAGAGGCTTGTGAAACAGAGAGATCTCATGTGCTGCAATTA 4470
Db 4021 GATGAGACTGTGTGTGTGAGAGGCTTGTGAAACAGAGAGATCTCATGTGCTGCAATTA 4080
QY 4471 CTGATCCCTCAAGCCGCACTAGCCGTAACCTCTGATGAGCCCGCAGTGTGCTGAC 4530
Db 4081 CTGATCCCTCAAGCCGCACTAGCCGTAACCTCTGATGAGCCCGCAGTGTGCTGAC 4140
QY 4531 TTTGTGAGCAAGTATGAGAGAGCTGATAGCCCGAGTGAACATCGGTGAGGTTGTCTTG 4590
Db 4141 TTTGTGAGCAAGTATGAGAGAGCTGATAGCCCGAGTGAACATCGGTGAGGTTGTCTTG 4200
```

QY	4591	GACAAATCGATGGACACAGGCTGTAGCCAGAGACAGTACAGAGAGGGGCTGGCTGAGAAC	4595
Db	4201	GACAAATCGATGGACACAGGCTGTAGCCAGAGACAGTACAGAGAGGGGCTGGCTGAGAAC	4260
QY	4651	ACGAGGCCCAACCGATGCGGAGCTGTTCAGCTTGAGCCAGTCTCTGGACCCGAAATGC	4710
Db	4261	ACGAGGCCCAACCGATGCGGAGCTGTTCAGCTTGAGCCAGTCTCTGGACCCGAAATGC	4320
QY	4711	AAAGATGGACTTACCAAGCCCTGAGAGAGACCCATCTCTCACTCATTTATGAACTCTGG	4770
Db	4321	AAAGATGGACTTACCAAGCCCTGAGAGAGACCCATCTCTCACTCATTTATGAACTCTGG	4380
QY	4771	GAGAGGGGCAAAAAAGGAGCTCTCGCAATCAGCAGCTGA	4812
Db	4381	GAGAGGGGCAAAAAAGGAGCTCTCTCGCAATCAGCAGCTGA	4422

QY 1303 GAGAGCTCTGTTCCACATGCCCCCTGGAAAAATGAGATTTTACCAAAAATTCACACAG 1362
 DB 781 GAGAGCTCTGTTCCACATGCCCCCTGGAAAAATGAGATTTTACCAAAAATTCACACAG 840
 QY 1363 CTGCTACTTCTACAAAGACCTCAACCCGAGAGCCAAAGATCCCTGGTCAAGAGAACTGG 1422
 DB 841 CTGCTACTTCTACAAAGACCTCAACCCGAGAGCCAAAGATCCCTGGTCAAGAGAACTGG 900
 QY 1423 CCTGATTTATGTGAGAGAGATCGAGACATTTATTTAGATTCAGAGATTTATTTGGCCCA 1482
 DB 901 CCTGATTTATGTGAGAGAGATCGAGACATTTATTTAGATTCAGAGATTTATTTGGCCCA 960
 QY 1483 GGGCTGTGATACCCAAAGAACTCCGATATGCTATCTGAGGGGGCTGTGGAATTTGGAAAG 1542
 DB 961 GGGCTGTGATACCCAAAGAACTCCGATATGCTATCTGAGGGGGCTGTGGAATTTGGAAAG 1020
 QY 1543 TCACAACCTGGCCAGGAGGTGAGAGAGAGCTGGGGGAGAGCCAGCTGTATGAGGAGACCG 1602
 DB 1021 TCACAACCTGGCCAGGAGGTGAGAGAGAGCTGGGGGAGAGCCAGCTGTATGAGGAGACCG 1080
 QY 1603 TTCCAGCATGCTCTTACTTCACTTCACTGCAAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1662
 DB 1081 TTCCAGCATGCTCTTACTTCACTTCACTGCAAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1140
 QY 1663 GCTGAGCTCATCGGAAAAAGTGGGACAGCCACTCCGGCTCCCATTTGACAGATCTGTCT 1722
 DB 1141 GCTGAGCTCATCGGAAAAAGTGGGACAGCCACTCCGGCTCCCATTTGACAGATCTGTCT 1200
 QY 1723 AGGCGCAGAGCGGGCTGCTTTCATCTCGATGCTGTGATGATGAGCAGAGATGGGCTTTGACG 1782
 DB 1201 AGGCGCAGAGCGGGCTGCTTTCATCTCGATGCTGTGATGATGAGCAGAGATGGGCTTTGACG 1260
 QY 1783 GAGCCAGATCTTGAAGCTGTCTGCACTGAGGCAAGCAGAGCCGAGGAGTGCATGCTG 1842
 DB 1261 GAGCCAGATCTTGAAGCTGTCTGCACTGAGGCAAGCAGAGCCGAGGAGTGCATGCTG 1320
 QY 1843 GGCAGTTTGTGGGGAAAACTATACCTCCGAGGCAATCCTTCTGATCACGGCTCGGACC 1902
 DB 1321 GGCAGTTTGTGGGGAAAACTATACCTCCGAGGCAATCCTTCTGATCACGGCTCGGACC 1380
 QY 1903 ACAGCTCTGAGAACCTCAATCTTCTTGGAGCAGAGCAGTGGGATGAGGCTTGGGG 1962
 DB 1381 ACAGCTCTGAGAACCTCAATCTTCTTGGAGCAGAGCAGTGGGATGAGGCTTGGGG 1440
 QY 1963 TTCTGTAGTCCAGAGAGAAATTTCTACAGATTTTCAAGATGAAAGCAAGCA 2022
 DB 1441 TTCTGTAGTCCAGAGAGAAATTTCTACAGATTTTCAAGATGAAAGCAAGCA 1500
 QY 2023 ATTAGAGCTTTTAGTTGGTCAATCAACCAAGAGCTCTGGGCTCTGTGTCTTGTGCC 2082
 DB 1501 ATTAGAGCTTTTAGTTGGTCAATCAACCAAGAGCTCTGGGCTCTGTGTCTTGTGCC 1560
 QY 2083 TGGGTCTCCGCTGGCTGCACTTGCCTGATGCAAGATGAAAGGAGAAAAATC 2142
 DB 1561 TGGGTCTCCGCTGGCTGCACTTGCCTGATGCAAGATGAAAGGAGAAAAATC 1620
 QY 2143 ACACTGATCTTCCAGACCAACCAACCTCTGTCTACATTAACCTTGCAGAGCTCTCAA 2202
 DB 1621 ACACTGATCTTCCAGACCAACCAACCTCTGTCTACATTAACCTTGCAGAGCTCTCAA 1680
 QY 2203 GCTCAGCATTGGGACCCCAAGCTCAGAGACCTCTGCTCTGCTGCTGAGGCAATGG 2262
 DB 1681 GCTCAGCATTGGGACCCCAAGCTCAGAGACCTCTGCTCTGCTGCTGAGGCAATGG 1740
 QY 2263 CAAAAAAGACCTTTTCACTGTCAGATGACCTCAGAGAGATGGTAAATGGGGCATC 2322
 DB 1741 CAAAAAAGACCTTTTCACTGTCAGATGACCTCAGAGAGATGGTAAATGGGGCATC 1800
 QY 2323 ATCTCCACCTTTTGAAGATGGGTATTTTCAAGAGACCCCAATCCCTCTGAGCTACAG 2382
 DB 1801 ATCTCCACCTTTTGAAGATGGGTATTTTCAAGAGACCCCAATCCCTCTGAGCTACAG 1860

QY 2383 TTCAATTCACCTCTGTTTCCAAAGTCTTTCAGAGATCTCTATGCTTTGAGAGATGAG 2442
 DB 1861 TTCAATTCACCTCTGTTTCCAAAGTCTTTCAGAGATCTCTATGCTTTGAGAGATGAG 1920
 QY 2443 AAGGGAGAGGTAAACATTTCTAATTGCAATCATAAGATTTGAAAGACGCTAGAGCATAT 2502
 DB 1921 AAGGGAGAGGTAAACATTTCTAATTGCAATCATAAGATTTGAAAGACGCTAGAGCATAT 1980
 QY 2503 GGAATACATGAGCTGTGTTGGGCAATCAACACAGTTTCTATTTGGGCTGTAAATGAT 2562
 DB 1981 GGAATACATGAGCTGTGTTGGGCAATCAACACAGTTTCTATTTGGGCTGTAAATGAT 2040
 QY 2563 GAGGGGAGAGAGATGAGAGAACTTTTCACTGCGGCTGTCTCAGGGAGAGAACTGG 2622
 DB 2041 GAGGGGAGAGAGATGAGAGAACTTTTCACTGCGGCTGTCTCAGGGAGAGAACTGG 2100
 QY 2623 ATGCAATGAGTCCCGTCCCTGAGCTGCTGCAAGCCCACTCTCTGAGATCCCTCAC 2682
 DB 2101 ATGCAATGAGTCCCGTCCCTGAGCTGCTGCAAGCCCACTCTCTGAGATCCCTCAC 2160
 QY 2683 TGCTGTATGAGACTTCGGAACAAACGTTCTTGACACAAAGTATGAGCCCATTTGAGAA 2742
 DB 2161 TGCTGTATGAGACTTCGGAACAAACGTTCTTGACACAAAGTATGAGCCCATTTGAGAA 2220
 QY 2743 ATGGGCAATGTGTAGAAAACAGACATGAGGCTTTAGTGTGACATTTTGCATTAATTC 2802
 DB 2221 ATGGGCAATGTGTAGAAAACAGACATGAGGCTTTAGTGTGACATTTTGCATTAATTC 2280
 QY 2803 ACCCGCACATGTAAGAAAGCTTCACTGATTTGAGGGCAGGACACAGATCAATGAGAC 2862
 DB 2281 ACCCGCACATGTAAGAAAGCTTCACTGATTTGAGGGCAGGACACAGATCAATGAGAC 2340
 QY 2863 CCCACATGTAAGTCTGTTCAAGTGGGCTCCAGTCAAGATGCTATTTGAGATTTCTC 2922
 DB 2341 CCCACATGTAAGTCTGTTCAAGTGGGCTCCAGTCAAGATGCTATTTGAGATTTCTC 2400
 QY 2923 TTCTCCGTCTTAAGTCAACAGAAACCTGAAAGAGTGTGACATTAAGTGAATTCGCTG 2982
 DB 2401 TTCTCCGTCTTAAGTCAACAGAAACCTGAAAGAGTGTGACATTAAGTGAATTCGCTG 2460
 QY 2983 ACCCACTCTGAGTAAAGTCTTTGTAAAGCCTTGAGAGCGCTCGCTGCTCTGAGAG 3042
 DB 2461 ACCCACTCTGAGTAAAGTCTTTGTAAAGCCTTGAGAGCGCTCGCTGCTCTGAGAG 2520
 QY 3043 ACCCTGCGGTGGCTGGCTGTGCTCAACGCTGAGAGCTGCAAGAGACCTTGCCTTTGG 3102
 DB 2521 ACCCTGCGGTGGCTGGCTGTGCTCAACGCTGAGAGCTGCAAGAGACCTTGCCTTTGG 2580
 QY 3103 CTGAGAGCCAAACCAACCTCTGACCGAGCTGACCTGAGCTTCAATGTGCTACAGATGCT 3162
 DB 2581 CTGAGAGCCAAACCAACCTCTGACCGAGCTGACCTGAGCTTCAATGTGCTACAGATGCT 2640
 QY 3163 GAGGCAAAACCTTTGCGAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACTGAG 3222
 DB 2641 GAGGCAAAACCTTTGCGAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACTGAG 2700
 QY 3223 CTGTGAGCTGTGCTCTACGTTGATCTGTGCTCAAGAGCTTGGCTTGTGTGCTG 3282
 DB 2701 CTGTGAGCTGTGCTCTACGTTGATCTGTGCTCAAGAGCTTGGCTTGTGTGCTG 2760
 QY 3283 AGCCCAAGCTTGAAGAGTGAACCTGAGAGAGAAACCACTGAGATGAGCTTGGCGTGA 3342
 DB 2761 AGCCCAAGCTTGAAGAGTGAACCTGAGAGAGAAACCACTGAGATGAGCTTGGCGTGA 2820
 QY 3343 CTGCTCTGAGAGGGCTCAGGATCTCTGCTCAAACTCATAGGCTGGGGCTGAGACAG 3402
 DB 2821 CTGCTCTGAGAGGGCTCAGGATCTCTGCTCAAACTCATAGGCT----- 2867
 QY 3403 ACAACTCTGATGATGATGAGGAGAGAACTGAGGGCCCTGAGAGAGAGAACTCTAG 3462
 DB 2868 ----- 2867
 QY 3463 CTGCTCATCTTGAAGAGAGAGAAACCAAGTGTGATGACCCCTCACTGAGGGCTGTGATAG 3522

```

Db      2868 -----GAAACCAAGTGATGACCCCTTCTAGAGGCCCTGGATACG 2907
QY      3523 GGAAGATGATGATATAGACATCTCTCACTCAAGCGGCAAGACTCCGATCAGAGAGGCG 3582
Db      2908 GGAGATGATGATATAGACATCTCTCACTCAAGCGGCAAGACTCCGATCAGAGAGGCG 2967
QY      3583 GCTTCCCATGTTGCTCAAGGCTATCTCAAACTCTGAGACGTGAGCAAGATCTTCCCAAT 3642
Db      2968 GCTTCCCATGTTGCTCAAGGCTATCTCAAACTCTGAGACGTGAGCAAGATCTTCCCAAT 3027
QY      3643 GCTGATATGACAGAGAAAGCTCCCAAGAGTAGTACCGGTGAACTCTTGCGTGCCT 3702
Db      3028 GCTGATATGACAGAGAAAGCTCCCAAGAGTAGTACCGGTGAACTCTTGCGTGCCT 3087
QY      3703 TCTCTGCTCTTCAGAGGGAAGCTGACATACGAGCTTTGGGGACTGACGATGACTTCTGG 3762
Db      3088 TCTCTGCTCTTCAGAGGGAAGCTGACATACGAGCTTTGGGGAAGCTGACGATGACTTCTGG 3147
QY      3763 GGGCCCAAGGGGCTGTGGTACTGAGAGTAGTGAAGCAAGAAAGAACTGTACCGAGTT 3822
Db      3148 GGGCCCAAGGGGCTGTGGTACTGAGAGTAGTGAAGCAAGAAAGAACTGTGTACCGAGTT 3207
QY      3823 CACTTCCCTGTAGCTGAGCTCTCACTGCGCTGAGCCCAACGAGGCTCTGCTTGTGATGAGA 3882
Db      3208 CACTTCCCTGTAGCTGAGCTCTCACTGCGCTGAGCCCAACGAGGCTCTGCTTGTGATGAGA 3267
QY      3883 GAAGCGGTGACCGTTGAGATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3942
Db      3268 GAAGCGGTGACCGTTGAGATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3327
QY      3943 CCACAGCAGCTGAGTGTGTGGGCAAGGCTCTGTGCACTCAAGGCTGAGCTGAGCT 4002
Db      3328 CCACAGCAGCTGAGTGTGTGGGCAAGGCTCTGTGCACTCAAGGCTGAGCTGAGCT 3387
QY      4003 GTGGAAGCTGTGACCTCCCTCACTTGTGTGGCTCTCCAAAGGGGCAAGTGTGACATCC 4062
Db      3388 GTGGAAGCTGTGACCTCCCTCACTTGTGTGGCTCTCCAAAGGGGCAAGTGTGACATCC 3447
QY      4063 CTGTTCCTCAATGGCCCACTTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCCAGCGGCTG 4122
Db      3448 CTGTTCCTCAATGGCCCACTTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCCAGCGGCTG 3507
QY      4123 GAGCTGATACATAGTGTCTGAGAAAGCCCAAGCTTCTCCCTTGGAGAGTCTCTGAAA 4182
Db      3508 GAGCTGATACATAGTGTCTGAGAAAGCCCAAGCTTCTCCCTTGGAGAGTCTCTGAAA 3567
QY      4183 ATGATCATATATGCCCTGAGCTTCACTCCGCTCACTCTGTGGGTGCTTTCACACCGC 4242
Db      3568 ATGATCATATATGCCCTGAGCTTCACTCCGCTCACTCTGTGGGTGCTTTCACACCGC 3627
QY      4243 GTCCATCTGAGAGAGTCACTTTCACCTCTACCTGATCCCAAGTGACTGCTCATTCGG 4302
Db      3628 GTCCATCTGAGAGAGTCACTTTCACCTCTACCTGATCCCAAGTGACTGCTCATTCGG- 3586
QY      4303 AAGGAATGAGAGCTGTCTATTCGAAGCCCTGAGAGAAACAAGCTTCTGAGATTAC 4362
Db      3687 --GGAATGAGAGCTGTCTATTCGAAGCCCTGAGAGAAACAAGCTTCTGAGATTAC 3744
QY      4363 GTTGGCACTTGGGATCAGGGATCAGGCTGACAGTGAAGAAACAAGAAAGTGAAGTCTG 4422
Db      3745 GTTGGCACTTGGGATCAGGGATCAGGCTGACAGTGAAGAAACAAGAAAGTGAAGTCTG 3804
QY      4423 GTGTGGAGAGCTTGTGTAACAGAGATCTCATGCTTCAACTACTGATCCCTCA 4482
Db      3805 GTGTGGAGAGCTTGTGTAACAGAGATCTCATGCTTCAACTACTGATCCCTCA 3864
QY      4483 GCGCCGATAGCCGTACTTCACTCTGAGAGCCCGGAGTGTGCTGCACTTTGTGAGACAG 4542
Db      3865 GCGCCGATAGCCGTACTTCACTCTGAGAGCCCGGAGTGTGCTGCACTTTGTGAGACAG 3924
QY      4543 TATGAGAGAGCTGATAGCCGAGTGAACATCGGTGAGGTTGCTTGAACAATCTCAT 4602

```

```

Db      3925 TATGAGAGAGCTGATAGCCGAGTGAATCGGTGTGGAGGTTCTTGAACAATGTCAT 3984
QY      4603 GGAAGAGTGTGAGCCAGAGACATTCAGAGAGGTTGCTGCTGAGAAACAGAGGCCAGC 4662
Db      3985 GGAAGAGTGTGAGCCAGAGACATTCAGAGAGGTTGCTGCTGAGAAACAGAGGCCAGC 4044
QY      4663 CAGATGCGGAAGTGTTCAGCTTGAAGCAGTCTCTGGGACCGGAAGTGAAGTGAAGTCTC 4722
Db      4045 CAGATGCGGAAGTGTTCAGCTTGAAGCAGTCTCTGGGACCGGAAGTGAAGTGAAGTCTC 4104
QY      4723 TACCAAGCCCTGAGAGAGACCCATCTCACTTATGAGAACTCTGGAGAGAGGCGAGC 4782
Db      4105 TACCAAGCCCTGAGAGAGACCCATCTCACTTATGAGAACTCTGGAGAGAGGCGAGC 4164
QY      4783 AAAAGGAGTCTCTGCACTGACAGCTGA 4812
Db      4165 AAAAGGAGTCTCTGCACTGACAGCTGA 4194

RESULT 8
AAd02762
ID AAd02762 standard; cDNA; 4329 BP.
XX
AC AAd02762;
XX
DT 31-MAY-2001 (first entry)
XX
DE Human NB-ARC and CARD containing protein (NAC) delta isoform cDNA.
XX
KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW cysteine aspartyl protease; apoptosis; cytokine production;
KW cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT 1..4329
FT /product= a
FT /delta isoform= "Human NB-ARC and CARD containing protein (NAC)"
FT 1..2868
FT misc_feature
FT /tag= b
FT /note= "Corresponds to 1-2868 residues of human NAC"
FT 2869..4329
FT /tag= c
FT /note= "Corresponds to 2962-4422 residues of human NAC"
FT beta isoform (AAd02760)"
XX
XX NO200116170-A2.
XX
XX PD 08-MAR-2001.
XX
XX PF 01-SEP-2000; 2000MO-US24152.
XX
XX PR 01-SEP-1999; 99US-0388221.
XX
XX (BURN-) BURHAM INST.
XX
XX PI Reed JC;
XX
XX DR WPI; 2001-183258/18.
XX DR P-PSDB; AAY72671.
XX
XX PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
XX PT domains, used to produce polypeptides for screening for modulators of
XX PT apoptosis -
XX
XX PS Claim 4; Page 148-154; 184pp; English.
XX
XX CC The present sequence is a human NB-ARC and CARD containing protein

```


CC (NAC) delta isoform cDNA. NAC delta isoform represents the NAC splice
 CC variant in which one of the splice region is absent in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signaling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukemias.

XX Sequence 4329 BP; 1018 A; 1202 C; 1208 G; 901 T; 0 other;

Query Match 72.6%; Score 3952; DB 22; Length 4329;

Best Local Similarity 94.9%; Pred. No. 0;

Matches 4197; Conservative 0; Mismatches 0; Indels 225; Gaps 2;

QY 523 ATGGCTGCGGAGACCTGGGCGCGCTGGCTGTACTTGAAGTTCTTGAAGAAGAGAG 582
 Db 1 ATGGCTGCGGAGACCTGGGCGCGCTGGCTGTACTTGAAGTTCTTGAAGAAGAGAG 60
 QY 583 CTGAAGAGTTTCCAGCTTCTGCTCCCAATTAAGCGCATCCAGAGAGCTTTGGGTAG 642
 Db 61 CTGAAGAGTTTCCAGCTTCTGCTCCCAATTAAGCGCATCCAGAGAGCTTTGGGTAG 120
 QY 643 ACACCCGCTCAGCCAGAGAGAAGAGATGAGATGAGAGTGGCTCTGTAAGTGGTCTAG 702
 Db 121 ACACCCGCTCAGCCAGAGAGAAGAGATGAGATGAGAGTGGCTCTGTAAGTGGTCTAG 180
 QY 703 TATGGGAGACAGCGGCGCTGGAGCTTACCTTCATACCTGGAGAGAGATGGGGCTGAG 762
 Db 181 TATGGGAGACAGCGGCGCTGGAGCTTACCTTCATACCTGGAGAGAGATGGGGCTGAG 240
 QY 763 TCACTGTGCGCCAGAGCCAGAGAGAGGAGAGAGCACTCCCTCATTCCTCCCTACAGCCCA 822
 Db 241 TCACTGTGCGCCAGAGCCAGAGAGAGGAGAGAGCACTCCCTCATTCCTCCCTACAGCCCA 300
 QY 823 AGTGAACCCACCTGGGGTCTCCCAAGCCAGACCACTCCACAGAGTGTAAATGACCTGG 882
 Db 301 AGTGAACCCACCTGGGGTCTCCCAAGCCAGACCACTCCACAGAGTGTAAATGACCTGG 360
 QY 883 ATTCATGAATTTGGCGGCGGGGTGACACCCAGGGCTCAGAGAGAAGGGTTTGAACAGCTG 942
 Db 361 ATTCATGAATTTGGCGGCGGGGTGACACCCAGGGCTCAGAGAGAAGGGTTTGAACAGCTG 420
 QY 943 CCTGACATCTGAGAGCGCGCTGGAGAGAAATCTGCTCACTCCTCTACCAAGCTCTT 1002
 Db 421 CCTGACATCTGAGAGCGCGCTGGAGAGAAATCTGCTCACTCCTCTACCAAGCTCTT 480
 QY 1003 CCNAGCTCCCAAGACCATGATCTCCNAGCCAGAGATCAACCAAGCCGCCCATCTACA 1062
 Db 481 CCNAGCTCCCAAGACCATGATCTCCNAGCCAGAGATCAACCAAGCCGCCCATCTACA 540
 QY 1063 GCAATGCTGGGAGGCTGGGGATCCCACTCAGCCCAAGCTTACAGACCCAGAGAGCAGAG 1122
 Db 541 GCAATGCTGGGAGGCTGGGGATCCCACTCAGCCCAAGCTTACAGACCCAGAGAGCAGAG 600
 QY 1123 GCTCCCTGGGAGACCAATGGGCTCTGATGAAAGTCAAGAAATTTACTACAGAAATCAAG 1182
 Db 601 GCTCCCTGGGAGACCAATGGGCTCTGATGAAAGTCAAGAAATTTACTACAGAAATCAAG 660
 QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGAGAGCCCAATGGGCAAGGGTGTAGAGAG 1242
 Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGAGAGCCCAATGGGCAAGGGTGTAGAGAG 720
 QY 1243 CCCCCCAGGCGGACACAGCTTACAGCCCAACCAACCATGGAGAGCTTTGTGAGA 1302
 Db 721 CCCCCCAGGCGGACACAGCTTACAGCCCAACCAACCATGGAGAGCTTTGTGAGA 780
 QY 1303 GAGAGCTCTGTTCACATGGGCTTGAAGAAATGAGGATTTTAAACAAAATTCACACAG 1362

Db 781 GAGAGCTCTGTTCACATGGGCTTGAAGAAATGAGATTTTAAACAAAATTCACACAG 840
 QY 1363 CTGCTACTTTTCAAAAGACTCAACCCAGAGCCAAAGATCCCTGGTCAAGAGAGCTGG 1422
 Db 841 CTGCTACTTTTCAAAAGACTCAACCCAGAGCCAAAGATCCCTGGTCAAGAGAGCTGG 900
 QY 1423 CCTGATATGTGAGAGAGATCGAGAGCATTTAATTGATCAGAGACTTATTGGCCCA 1482
 Db 901 CCTGATATGTGAGAGAGATCGAGAGCATTTAATTGATCAGAGACTTATTGGCCCA 960
 QY 1483 GGGCTGAGTACCCAAAGACCTCGATACATACCTGAGAGGGGGCTGAGATTGGGAG 1542
 Db 961 GGGCTGAGTACCCAAAGACCTCGATACATACCTGAGAGGGGGCTGAGATTGGGAG 1020
 QY 1543 TCAACATGCGCCAGAGCAGATGAGAGAGCCTGGGGAGAGAGCCAGCTGATAGGAGCCGC 1602
 Db 1021 TCAACATGCGCCAGAGCAGATGAGAGAGCCTGGGGAGAGAGCCAGCTGATAGGAGCCGC 1080
 QY 1603 TTCCAGCATGTCTTTACTTCACTGACAGAGAGCTGGCCAGTCCAGAGTGGTGAATCTC 1662
 Db 1081 TTCCAGCATGTCTTTACTTCACTGACAGAGAGCTGGCCAGTCCAGAGTGGTGAATCTC 1140
 QY 1663 GCTGAGCTCATGGAAGAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCCGTCT 1722
 Db 1141 GCTGAGCTCATGGAAGAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCCGTCT 1200
 QY 1723 AGGCCAGAGCGGCTGCTCTTCACTCTCGATGGTATGATGAGCCAGATGGGTCTTGGAG 1782
 Db 1201 AGGCCAGAGCGGCTGCTCTTCACTCTCGATGGTATGATGAGCCAGATGGGTCTTGGAG 1260
 QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGACCTGAGCCAGCCAGAGCCGGCGGATGACATGCTG 1842
 Db 1261 GAGCCGAGTTCTGAGCTCTGTCTGACCTGAGCCAGCCAGAGCCGGCGGATGACATGCTG 1320
 QY 1843 GGCAGTTTGTGGGGGAAAATTAATCTCCAGAGGATCTTCTGATCAAGGCTGGAGCC 1902
 Db 1321 GGCAGTTTGTGGGGGAAAATTAATCTCCAGAGGATCTTCTGATCAAGGCTGGAGCC 1380
 QY 1903 ACAGCTCTGAGAGACCTGATCTCTTTTGGAGCAGGACGTTGGGTAGAGGCTCTGGGG 1962
 Db 1381 ACAGCTCTGAGAGACCTGATCTCTTTTGGAGCAGGACGTTGGGTAGAGGCTCTGGGG 1440
 QY 1963 TTCTCTGAGTCCAGAGAGAGAAATTTTCAACATATTTCAACAGATGAAGGCAAGA 2022
 Db 1441 TTCTCTGAGTCCAGAGAGAGAAATTTTCAACATATTTCAACAGATGAAGGCAAGA 1500
 QY 2023 ATTAGAGCTTTAGGTTGGTCAATCAACAAAGAGCTCTGGGCGCTGTGTCTTGTGCC 2082
 Db 1501 ATTAGAGCTTTAGGTTGGTCAATCAACAAAGAGCTCTGGGCGCTGTGTCTTGTGCC 1560
 QY 2083 TGGGTGTCTGTGGCTGCACTTGCTGATGAGCAGATGAAAGGAGAAAACCTC 2142
 Db 1561 TGGGTGTCTGTGGCTGCACTTGCTGATGAGCAGATGAAAGGAGAAAACCTC 1620
 QY 2143 ACATGATCTTCCAAACCAACCAACCTCTGTCTTCAATTAACCTTCCAGGCTCTCCAA 2202
 Db 1621 ACATGATCTTCCAAACCAACCAACCTCTGTCTTCAATTAACCTTCCAGGCTCTCCAA 1680
 QY 2203 GCTCAGCATTTGGAGCCCAAGCTCAGAGACCTCTGCTCTGCTGCTGAGAGGCAATCTGG 2262
 Db 1681 GCTCAGCATTTGGAGCCCAAGCTCAGAGACCTCTGCTCTGCTGCTGAGAGGCAATCTGG 1740
 QY 2263 CAAAAAAGACCTTTTCAATGATCAGATGACCTCAGAGAGCATGGGTGATGGGCGATC 2322
 Db 1741 CAAAAAAGACCTTTTCAATGATCAGATGACCTCAGAGAGCATGGGTGATGGGCGATC 1800
 QY 2323 ATCTCAGCTTTTGAAGTGGGTATCTTCAAGAGCAACCCATCCCTGAGCTTACAGC 2382
 Db 1801 ATCTCAGCTTTTGAAGTGGGTATCTTCAAGAGCAACCCATCCCTGAGCTTACAGC 1860
 QY 2383 TTCAATTCACCTCTGTTCCAAGAGTCTTGTGAGCAATGCTTATGTCTTGAAGATGAG 2442
 Db 1861 TTCAATTCACCTCTGTTCCAAGAGTCTTGTGAGCAATGCTTATGTCTTGAAGATGAG 1920

QY 2443 AAGGAGAGGTAATTAATTCATCATAGATTTGAAAAGAGCTGAAAGCATAT 2502
 DB 1921 AAGGGAGAGGTAATTAATTCATCATAGATTTGAAAAGAGCTGAAAGCATAT 1980
 QY 2503 GGAATACATGAGCTGTGTTGGGCAATCAACAACGTTTCTATTGGGCTTTAAGTAT 2562
 DB 1981 GGAATACATGAGCTGTGTTGGGCAATCAACAACGTTTCTATTGGGCTTTAAGTAT 2040
 QY 2563 GAGGGGAGAGAGATGAGAAATCTTTCACTGCGGCTGTCTCAGGGAGAACTGT 2622
 DB 2041 GAGGGGAGAGAGATGAGAAATCTTTCACTGCGGCTGTCTCAGGGAGAACTGT 2100
 QY 2623 ATGAGTGGGTCCGCTCCCTGAGCTGTGCTGAGCACTCTCTGAGATCCCTCCAC 2682
 DB 2101 ATGAGTGGGTCCGCTCCCTGAGCTGTGCTGAGCACTCTCTGAGATCCCTCCAC 2160
 QY 2683 TGCTTTGACAGAGACTCGGACAAAACGTTCCCTGACACAGATGATGAGCCATTTGAAAGAA 2742
 DB 2161 TGCTTTGACAGAGACTCGGACAAAACGTTCCCTGACACAGATGATGAGCCATTTGAAAGAA 2220
 QY 2743 ATGGGCAATGTGTGTAGAAACAGACATGAGACTCTTAATGTGCACTTTCTGATTAATTC 2802
 DB 2221 ATGGGCAATGTGTGTAGAAACAGACATGAGACTCTTAATGTGCACTTTCTGATTAATTC 2280
 QY 2803 AGCGGCAACGTGAAGAGCTTCACTGATGAGGGGAGGAGCAAGATCAACATGAGC 2862
 DB 2281 AGCGGCAACGTGAAGAGCTTCACTGATGAGGGGAGGAGCAAGATCAACATGAGC 2340
 QY 2863 CCCACATGTGATGCTGTTGAGGTGGGTCCAGTCAAGATGCTATTGAGCAATTC 2922
 DB 2341 CCCACATGTGATGCTGTTGAGGTGGGTCCAGTCAAGATGCTATTGAGCAATTC 2400
 QY 2923 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGAACTTAATGTGAAATTCGCTG 2982
 DB 2401 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGAACTTAATGTGAAATTCGCTG 2460
 QY 2983 AGCACTCTGAGTGAAGAGCTTGTGTAGACCTGAGAGCGCTCGCTGCTCTGAG 3042
 DB 2461 AGCACTCTGAGTGAAGAGCTTGTGTAGACCTGAGAGCGCTCGCTGCTCTGAG 2520
 QY 3043 ACCCTGCGGTGTGCTGTGCTGCTCAAGCTGAGAGACTGCAAGAACTTTCCTTTGGG 3102
 DB 2521 ACCCTGCGGTGTGCTGTGCTGCTCAAGCTGAGAGACTGCAAGAACTTTCCTTTGGG 2580
 QY 3103 CTGAGAGCAACCAAGACCTGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 3162
 DB 2581 CTGAGAGCAACCAAGACCTGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 2640
 QY 3163 GGAGCCAAACCTTTGCTGAGAGCTGAGAGCGGAGCTGAGCTTCAAGCTGAG 3222
 DB 2641 GGAGCCAAACCTTTGCTGAGAGCTGAGAGCGGAGCTGAGCTTCAAGCTGAG 2700
 QY 3223 CTGCTGAGCTGTGCTGCTCAAGCTGAGCTGAGAGCTGAGCTTCTGTAGTCC 3282
 DB 2701 CTGCTGAGCTGTGCTGCTCAAGCTGAGCTGAGAGCTGAGCTTCTGTAGTCC 2760
 QY 3283 AGCCCAAGCTGAGAGAGCTGAGCTGAGAGCAACCTGATGAGCTTGGGCTGGA 3342
 DB 2761 AGCCCAAGCTGAGAGAGCTGAGCTGAGAGCAACCTGATGAGCTTGGGCTGGA 2820
 QY 3343 CTGCTGAGTGGGGCTGAGAGCACTGAGCTGAGAGCAACCTGAGGGGCTGAGCAAG 3402
 DB 2821 CTGCTGAGTGGGGCTGAGAGCACTGAGCTGAGAGCAACCTGAGGGGCTGAGCAAG 2867
 QY 3403 ACAACTGAGTGTGATGAGTGAAGGAGCAAGCTGAGGGGCTGAGAGCAAGAACTGAG 3462
 DB 2868 ----- 2867
 QY 3463 CTGCTCATCTTACAGACAGGAAACCAAGTGTATGATCCCTTACTGAGGGCTTGGATTCG 3522
 DB 2868 -----GAAACCAAGTGTATGATCCCTTACTGAGGGCTTGGATTCG 2907

QY 3523 GAGAGATGAGTATATGACATCTCTCACTCAAGGGGAGAGACTCGATCAGAGGGCG 3582
 DB 2908 GAGAGATGAGTATATGACATCTCTCACTCAAGGGGAGAGACTCGATCAGAGGGCG 2967
 QY 3583 GCTTCCCATGTGCTCAGGCTAATCTCAAACTCTGAGACGTGAGCAAGATCTTCCAAAT 3642
 DB 2968 GCTTCCCATGTGCTCAGGCTAATCTCAAACTCTGAGACGTGAGCAAGATCTTCCAAAT 3027
 QY 3643 GCTGAGATGAGAGAGAAAGTCCCAAGAGTATACCGGTGAGAACTTGTGTGCTGCT 3702
 DB 3028 GCTGAGATGAGAGAGAAAGTCCCAAGAGTATACCGGTGAGAACTTGTGTGCTGCT 3087
 QY 3703 TCTCTGCTCTCAAGGGGAGCTGATACAGAAAGCTTTGGGAGCTGAGATGCTTCTG 3762
 DB 3088 TCTCTGCTCTCAAGGGGAGCTGATACAGAAAGCTTTGGGAGCTGAGATGCTTCTG 3147
 QY 3763 GGCCCAAGGGGCTGTGCTACTGAGGTATGACAAAGAAAGAACTTGTACCGAGT 3822
 DB 3148 GGCCCAAGGGGCTGTGCTACTGAGGTATGACAAAGAAAGAACTTGTACCGAGT 3207
 QY 3823 CACTTCCCTGTAGTGGCTCTTACCGCTGACCAACAGGCTCTGTGTTGTATGAGA 3882
 DB 3208 CACTTCCCTGTAGTGGCTCTTACCGCTGACCAACAGGCTCTGTGTTGTATGAGA 3267
 QY 3883 GAAGCGGTGACCGGTGAGTTGAATTCGTGTGTGAGGACCAAGTCCCTGGGTGAGATCAC 3942
 DB 3267 GAAGCGGTGACCGGTGAGTTGAATTCGTGTGTGAGGACCAAGTCCCTGGGTGAGATCAC 3327
 QY 3943 CCACAGACAGCTGATGATGTGAGAGGGCTCTGCTGAGCAATCAAGGCTGAGCTGAGCT 4002
 DB 3328 CCACAGACAGCTGATGATGTGAGAGGGCTCTGCTGAGCAATCAAGGCTGAGCTGAGCT 3387
 QY 4003 GTGGAAGCTGTGACCTTCCCTCACTTTGTGCTCTTCCAAAGGGGCTATGTGACATCC 4062
 DB 3388 GTGGAAGCTGTGACCTTCCCTCACTTTGTGCTCTTCCAAAGGGGCTATGTGACATCC 3447
 QY 4063 CTGTTCCAAATGTGCGCACTTTAAAGAGAGGGAGTGCCTCGAGAAAGCCAGAGGCTG 4122
 DB 3448 CTGTTCCAAATGTGCGCACTTTAAAGAGAGGGAGTGCCTCGAGAAAGCCAGAGGCTG 3507
 QY 4123 GAGCTGATCATCATATGTTCTGAGAAACCCCAAGCTTCTCCCTTTGGAGTCCCTGAAA 4182
 DB 3508 GAGCTGATCATCATATGTTCTGAGAAACCCCAAGCTTCTCCCTTTGGAGTCCCTGAAA 3567
 QY 4183 ATGATCATATATGCTGCGCTTCAATTCCTGTCACCTCTGTGTGTGCTTTACACCGC 4242
 DB 3568 ATGATCATATATGCTGCGCTTCAATTCCTGTCACCTCTGTGTGTGCTTTACACCGC 3627
 QY 4243 GTCCATCTGAGAGAGTCACTTCCACTTCACTGATCCCAAGTGACTGCTCATTCG 4302
 DB 3628 GTCCATCTGAGAGAGTCACTTCCACTTCACTGATCCCAAGTGACTGCTCATTCG 3687
 QY 4303 ----- 4302
 DB 3688 AAGGCAATGATGATCTAGAAATGAATTCAGTTTGGGAATCCAAGCCACCCCG 3747
 QY 4303 ----- 4302
 DB 3748 CTGACCCCACTTATATGAGCTGTGTTACATGTGTCTGGGTTCTGATTCAGGATGCTG 3807
 QY 4303 -----AAGAACTGAGCTGTGCTATGAGAGCCCTGAGAAAGCAAGCTGTC 4350
 DB 3808 GAAATATCTCCCAAGAGAGTGAAGCTGTGCTATGAGAGCCCTGAGAAAGCAAGCTGTC 3867
 QY 4351 TCGGAGTTCTACGTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGCAAGAAA 4410
 DB 3868 TCGGAGTTCTACGTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGCAAGAAA 3927
 QY 4411 GATGAGACTCTGTGTGTGAGAGGCTTGTGTGAACCAAGAGATCTCATGCTTGCACATCT 4470
 DB 3928 GATGAGACTCTGTGTGTGAGAGGCTTGTGTGAACCAAGAGATCTCATGCTTGCACATCT 3987
 QY 4471 CTGATCCCTCAGGCCGATAGCGGTACCTTCACTCTGATGAGCCCGCAGTTGCTGCAC 4530

Db 3988 CTATCCCTCCAGCCGATACGCTTCACTTCACTTGGATGCCGCGAGTTGGCAC 4047
 QY 4531 TTTGTGACCAATATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4590
 Db 4048 TTTGTGACCAATATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4107
 QY 4591 GAAATATGATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4650
 Db 4108 GAAATATGATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4167
 QY 4651 ACAGAGCCGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4710
 Db 4168 ACAGAGCCGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4227
 QY 4711 AAGATGATGATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4770
 Db 4228 AAGATGATGATGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4287
 QY 4771 GAGAGAGAGAGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4812
 Db 4288 GAGAGAGAGAGAGAGAGCTGATAGCCGAGTGAATCGATGAGTTGCTTG 4329

RESULT 9

AAD02764 standard, cDNA; 4556 BP.

AAD02764;

31-MAY-2001 (first entry)

Human NAC beta isoform-CARD-X1 chimeric cDNA.

KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cytosine receptor signaling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukemia;
 KW chimeric protein; ss.

OS Homo sapiens.

Location/Qualifiers

FT CDS 1..4365 /tag= a /product= "Human NAC beta isoform-CARD-X1 chimera"

FT misc_feature 1..3234 /tag= b /note= "Corresponds to 1-3234 residues of human"

FT misc_feature 3235..4362 /tag= c /note= "Corresponds to 166-1293 residues of human"

FT CARD-X (AAD02763)"

PN W0200116170-A2.

PD 08-MAR-2001.

PF 01-SEP-2000; 2000WO-US24152.

PR 01-SEP-1999; 99US-0388221.

PA (BURN-) BURHAM INST.

PI Reed JC;

XX WPI; 2001-183258/18.

DR P-PSDB; AAY72673.

XX Novel nucleic acid encoding NB-ARC and caspase associated recruitment

PT domains, used to produce polypeptides for screening for modulators of

PT apoptosis -
 PS Disclosure; Page 162-168; 184p; English.
 CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
 CC beta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
 CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
 CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
 CC domain. CARD-X protein comprises a caspase-associated recruitment domain
 CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
 CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
 CC for screening modulators that modulates apoptosis, cytokine production,
 CC cytokine receptor signaling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory
 CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukemias.

SQ Sequence 4556 BP; 1100 A; 1223 C; 1256 G; 977 T; 0 other;

Query Match 62.5%; Score 3400.4; DB 22; Length 4556;

Best Local Similarity 93.7%; Pred. No. 0; Mismatches 236; Indels 3; Gaps 1;

QY 523 ATGGCTGGGAGAGCTGGGCGCCCTGGCTGATTTGAGATTCTGAAAGAGAGAG 582
 Db 1 ATGGCTGGGAGAGCTGGGCGCCCTGGCTGATTTGAGATTCTGAAAGAGAGAG 60
 QY 583 CTGAAGAGATTCAGCTTCTGCTGCGCAATTAAGSCATCCAGAGAGCTTTCGGGTAG 642
 Db 61 CTGAAGAGATTCAGCTTCTGCTGCGCAATTAAGSCATCCAGAGAGCTTTCGGGTAG 120
 QY 643 ACAACCGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
 Db 121 ACAACCGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 703 TATGGGAG 762
 Db 181 TATGGGAG 240
 QY 763 TCATGCTGGCCCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
 Db 241 TCATGCTGGCCCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 823 AGTGAACCCCACTGGGATCTCCAGCCCAACCACTCCAGCGAGTGTATGCCCTGG 882
 Db 301 AGTGAACCCCACTGGGATCTCCAGCCCAACCACTCCAGCGAGTGTATGCCCTGG 360
 QY 883 ATCCATGAATTCGCGGCGGGGTGACCCAGAGGCTCAGAGAGAGAGAGAGAGAG 942
 Db 361 ATCCATGAATTCGCGGCGGGGTGACCCAGAGGCTCAGAGAGAGAGAGAGAGAG 420
 QY 943 CCTGACATCTGAGAGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1002
 Db 421 CCTGACATCTGAGAGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 1003 CCAAGTCCCAAGACCTAGATCTCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
 Db 481 CCAAGTCCCAAGACCTAGATCTCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 1063 GCAGTCTGGGAG 1122
 Db 541 GCAGTCTGGGAG 600
 QY 1123 GCTCCTGGAG 1182
 Db 601 GCTCCTGGAG 660
 QY 1183 GAAAG 1242
 Db 661 GAAAG 720
 QY 1243 CCCCAG 1302

D	b	721	CCCCCAGGGGCAACACGAGCTTACAGCCCAACACCAAGGAGACCTTCTGTGAA	780
O	y	1303	GAGAGCCTCTGTTCCACATGAGCCCTGGAAAAATGAGATTTTAAACAAAAATTCA	1362
D	b	781	GAGAGCCTCTGTTCCACATGAGCCCTGGAAAAATGAGATTTTAAACAAAAATTCA	840
O	y	1363	CTGCTACTTCTTAACAAAGCTCTACCCCAAGAACCAAGATCCCTGTGTCAAGAA	1422
D	b	841	CTGCTACTTCTTAACAAAGCTCTACCCCAAGAACCAAGATCCCTGTGTCAAGAA	900
O	y	1423	CTGCTACTTCTTAACAAAGCTCTACCCCAAGAACCAAGATCCCTGTGTCAAGAA	1482
D	b	901	CTGCTACTTCTTAACAAAGCTCTACCCCAAGAACCAAGATCCCTGTGTCAAGAA	960
O	y	1483	GGCCTGATACCCCAAGAACTCGCATAGTCACTGACAGGGGCTGTGAAATTGGAA	1542
D	b	961	GGCCTGATACCCCAAGAACTCGCATAGTCACTGACAGGGGCTGTGAAATTGGAA	1020
O	y	1543	TCACACTGAGCCAGGCAAGTGAAGAGCCTGGGGGAGAGGCGACGTGATGGGAG	1602
D	b	1021	TCACACTGAGCCAGGCAAGTGAAGAGCCTGGGGGAGAGGCGACGTGATGGGAG	1080
O	y	1603	TTCCAGCATGCTTCTTACTTTCAGCTGACAGAGCTGGCCCACTCAAGTGTGA	1662
D	b	1081	TTCCAGCATGCTTCTTACTTTCAGCTGACAGAGCTGGCCCACTCAAGTGTGA	1140
O	y	1663	GCTGAGCTCATGCGAAAGATGGGACAGACCACTCCGGCTCCCATTTAGACAGAT	1722
D	b	1141	GCTGAGCTCATGCGAAAGATGGGACAGACCACTCCGGCTCCCATTTAGACAGAT	1200
O	y	1723	AGGCGACAGCGAGCTCTCTTCTTCTCGATGATGATGAGCCAGATGGGATCTTG	1782
D	b	1201	AGGCGACAGCGAGCTCTCTTCTTCTCGATGATGATGAGCCAGATGGGATCTTG	1260
O	y	1783	GAGCCGAGTCTGAGCTCTGTCTGCACTGAGCCACGACCGGCGGATGCACTGTG	1842
D	b	1261	GAGCCGAGTCTGAGCTCTGTCTGCACTGAGCCACGACCGGCGGATGCACTGTG	1320
O	y	1843	GGCAGTTGCTGGGGGAAACTATCTCCGAGGACCTTCTGATACAGGCTCGAAC	1902
D	b	1321	GGCAGTTGCTGGGGGAAACTATCTCCGAGGACCTTCTGATACAGGCTCGAAC	1380
O	y	1903	ACAGCTCTGACAAACCTCATCTTCTTGGAGCAGGACGTTGGGTAGAGTCTGTGG	1962
D	b	1381	ACAGCTCTGACAAACCTCATCTTCTTGGAGCAGGACGTTGGGTAGAGTCTGTGG	1440
O	y	1963	TTCTCTGAGTCCAGCAGAGAGGAAATTTTTCACAGATATTTTTCACAGATGA	2022
D	b	1441	TTCTCTGAGTCCAGCAGAGAGGAAATTTTTCACAGATATTTTTCACAGATGA	1500
O	y	2023	ATTAGAGCTTTAGGTGTGTAATCAACAAAGAGCTGTGGGCTGTGTCTGTGGCC	2082
D	b	1501	ATTAGAGCTTTAGGTGTGTAATCAACAAAGAGCTGTGGGCTGTGTCTGTGGCC	1560
O	y	2083	TGGGTCTCTGGCTGGCTGCACTTGTCTGATGACAGATGAAAGCGAAAGAAATC	2142
D	b	1561	TGGGTCTCTGGCTGGCTGCACTTGTCTGATGACAGATGAAAGCGAAAGAAATC	1620
O	y	2143	ACACTGATTTCCAAAGACCAACCAACCTCTGTCTTACATTAACCTTGCCAGGCTT	2202
D	b	1621	ACACTGATTTCCAAAGACCAACCAACCTCTGTCTTACATTAACCTTGCCAGGCTT	1680
O	y	2203	GCTACGCAATTGGGAGCCCAAGCTCAGAGACCTCTGTGCTGTGAGGGGCACTGG	2262
D	b	1681	GCTACGCAATTGGGAGCCCAAGCTCAGAGACCTCTGTGCTGTGAGGGGCACTGG	1740
O	y	2263	CAAAAAAAGACCTTTTCAAGTCAAGATACCTCAGAAAGCATGGTTAGATGGGCA	2322
D	b	1741	CAAAAAAAGACCTTTTCAAGTCAAGATACCTCAGAAAGCATGGTTAGATGGGCA	1800
O	y	2323	ATCTCACCTTCTTGAAGATGGGTATTTTCAAGACACCCCATCTCTGAGCTAC	2382

QY 3463 CTGCTCATTTTCAAGCAGAGGAAACCAAGTGTGATGACCCCTTACTGAGGGCTCGATACG 3522
 DB 2941 CTGCTCATTTTCAAGCAGAGGAAACCAAGTGTGATGACCCCTTACTGAGGGCTCGATACG 3000
 QY 3523 GGAGAGATGATATATGACATCTTCACTCAAGCGGAGAGACTGGGATCAGAGGGGG 3582
 DB 3001 GGAGAGATGATATATGACATCTTCACTCAAGCGGAGAGACTGGGATCAGAGGGGG 3060
 QY 3583 GCTTCCCATGTTGCTCAGGCTATCTCAACTCTGAGCGTACAGAGATCTTCCCAAT 3642
 DB 3061 GCTTCCCATGTTGCTCAGGCTATCTCAACTCTGAGCGTACAGAGATCTTCCCAAT 3120
 QY 3643 GCTGAGATTGCAAGAGAAAGCTTCCCAAGAGTATCCGGTGAACCTTTGCGTGGCT 3702
 DB 3121 GCTGAGATTGCAAGAGAAAGCTTCCCAAGAGTATCCGGTGAACCTTTGCGTGGCT 3180
 QY 3703 TCTCCGCTCTCAAGGGGAGCTGCACTAGAGGCTTTGGGGAGCTGAGCATCTTCTGG 3762
 DB 3181 TCTCCGCTCTCAAGGGGAGCTGCACTAGAGGCTTTGGGGAGCTGAGCATCTTCTGG 3240
 QY 3763 GGCCCCACGGGAGCTGTGAGTCTACTGAGTATGACAAAGAAAGAACTTGTACCGAGTT 3822
 DB 3241 GGCGCTGAAAGAAAGTGTGATGTTGATGATGATGATGATGATGATGATGATGATGAT 3300
 QY 3823 CACTTCCCTGATGAGTGGCTCTTACCGCTGGCCCAACAGGGTCTCTTGTGATGAGA 3882
 DB 3301 TGGTTCCTCCACAGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3360
 QY 3883 GAAAGGGTGAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3939
 DB 3361 GATGAGGTGACAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
 QY 3940 AACCCACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3999
 DB 3421 CAGCAGCAGTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
 QY 4000 GCTGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4059
 DB 3481 GAGGCTGTGAGCAGTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3540
 QY 4060 TCCCTGTTCCAAATGGGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4119
 DB 3541 TCCCTGTTCTGTTGCTGATTTAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
 QY 4120 GTGAGCTGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4179
 DB 3601 GTGAGCTGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
 QY 4180 AAAATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4239
 DB 3661 CGAGTGTGAG 3720
 QY 4240 CGGCTGATCAG 4299
 DB 3721 CACCCCTACCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
 QY 4300 CGAAG 4313
 DB 3781 ACAAG 3794
 RESULT 10
 AAD02765
 ID AAD02765 standard; cDNA; 4466 BP.
 XX
 AC AAD02765;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Human NAC gamma or delta isoform-CARD-X1 chimeric cDNA.
 XX
 KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;

KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signaling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukemia;
 KW chimeric protein; ss.
 OS
 XX Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 1..4275
 FT /tag= a
 FT /product= "Human NAC gamma/delta isoform-CARD-X1 chimera"
 FT misc_feature
 FT 1..3144
 FT /tag= b
 FT /note= "Corresponds to 1-3144 residues of human
 FT NAC gamma (AAD02761) or delta (AAD02762) isoform"
 FT 3145..4272
 FT /tag= c
 FT /note= "Corresponds to 166-1293 residues of human
 FT CARD-X (AAD02763)"
 FT
 PN
 XX WO200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 PA (BURN-) BURHAM INST.
 XX
 PI Reed JC;
 XX
 DR WPI: 2001-183258/18.
 XX
 DR P-PSDB; AAY72674.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domain, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 XX
 PS Disclosure; Page 173-178; 184pp; English.
 XX
 CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
 CC gamma or delta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
 CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
 CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
 CC domain. CARD-X protein comprises a caspase-associated recruitment domain
 CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
 CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
 CC for screening modulators that modulates apoptosis, cytokine production,
 CC cytokine receptor signaling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory
 CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukemias.
 CC
 XX
 SQ Sequence 4466 BP; 1075 A; 1201 C; 1227 G; 963 T; 0 other;
 Query Match 59.0%; Score 3210.4; DB 22; Length 4466;
 Best Local Similarity 91.3%; Pred. No. 0;
 Matches 3465; Conservative 0; Mismatches 236; Indels 93; Gaps 2;
 QY 523 ATGGCTGCGAGAGCTGTGGGCGCGCTGTGCTTACTTGAAGTCTTGAAGAGAGAGAG 582
 DB 1 ATGGCTGCGAGAGCTGTGGGCGCGCTGTGCTTACTTGAAGTCTTGAAGAGAGAGAG 60
 QY 583 CTGAAGAGATCCAGCTTCTGTGCGCAATTAAGCGCATCCAGAGCTCTTGGGTGAG 642
 DB 61 CTGAAGAGATCCAGCTTCTGTGCGCAATTAAGCGCATCCAGAGCTCTTGGGTGAG 120
 QY 643 ACACCCGCTCAGCCAG 702
 DB 121 ACACCCGCTCAGCCAG 180

QY 703 TATGGGAGAGAGGGGCTTGGGACTTAGCCCTCCATACCTTGGAGAGAGATGGGGCTTAGG 762
DB 181 TATGGGAGAGAGGGGCTTGGGACTTAGCCCTCCATACCTTGGAGAGAGATGGGGCTTAGG 240
QY 763 TCACTGAGCCCAAGGCCAGAGAAAGGGGCAAGGCACTCTCCCTCATTTCCCTACAGGCCA 822
DB 241 TCACTGAGCCCAAGGCCAGAGAAAGGGGCAAGGCACTCTCCCTCATTTCCCTACAGGCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTTCAACCGGAGTCTAATGCCCTGG 882
DB 301 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTTCAACCGGAGTCTAATGCCCTGG 360
QY 883 ATTCATGAATTTGCCGGGGGGGTGCAACCAAGGCTCAAGAGAAAGGTTTGAACAGCTG 942
DB 361 ATTCATGAATTTGCCGGGGGGGTGCAACCAAGGCTCAAGAGAAAGGTTTGAACAGCTG 420
QY 943 CTGACACATCTGGAAGCCGCTGGAGAGAAATCTGTGCTCACTCTCTCAACAGCTCTT 1002
DB 421 CTGACACATCTGGAAGCCGCTGGAGAGAAATCTGTGCTCACTCTCTCAACAGCTCTT 480
QY 1003 CCAAGCTCCCAAGCATGAGTCTCAAGCCAGAGTCAACCAACGCCCCCAATCCACA 1062
DB 481 CCAAGCTCCCAAGCATGAGTCTCAAGCCAGAGTCAACCAACGCCCCCAATCCACA 540
QY 1063 GCAGTGCTGGGAGCTGGGGATCCCACTTCAAGCCAGCTTAGACCCAGAGAGAGAG 1122
DB 541 GCAGTGCTGGGAGCTGGGGATCCCACTTCAAGCCAGCTTAGACCCAGAGAGAGAG 600
QY 1123 GCATCTGGGAGCCCAATGGCCCTGTGATGAACGTCAGAAATTTACTACAGAAATCAGA 1182
DB 601 GCATCTGGGAGCCCAATGGCCCTGTGATGAACGTCAGAAATTTACTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCATAGGGGAGGGTGTAGAACG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCATAGGGGAGGGTGTAGAACG 720
QY 1243 CCCCCAAGGCGCACACAGCTTACAGCCCAACCAACCTAGGAGGCTTCTGTAGA 1302
DB 721 CCCCCAAGGCGCACACAGCTTACAGCCCAACCAACCTAGGAGGCTTCTGTAGA 780
QY 1303 GAGAGCTCTGTTCCATGAGGCTCTGGAAAAAATGAGATTTTAAACAAAAATTCACAG 1362
DB 781 GAGAGCTCTGTTCCATGAGGCTCTGGAAAAAATGAGATTTTAAACAAAAATTCACAG 840
QY 1363 CTGCTACTTTCACAAAGACTCAACCCAGAGCCAGATCCCTGTCAAGAGAGCTGG 1422
DB 841 CTGCTACTTTCACAAAGACTCAACCCAGAGCCAGATCCCTGTCAAGAGAGCTGG 900
QY 1423 CTTGATTTATGTGAGAGAGATCGAGACATTTAATTGAGTCAAGACATTTATTTGGCCCA 1482
DB 901 CTTGATTTATGTGAGAGAGATCGAGACATTTAATTGAGTCAAGACATTTATTTGGCCCA 960
QY 1483 GAGCTGATTAACCAAGAACTCGCATGTCTACTGAGAGGGGGGCTGTGAGAAATTTGGGAG 1542
DB 961 GAGCTGATTAACCAAGAACTCGCATGTCTACTGAGAGGGGGGCTGTGAGAAATTTGGGAG 1020
QY 1543 TCAACACTGGCCAGGCGAGTGAAGAAAGCTTGGGGAGAGAGCCAGCTGTATGGGGAGCCG 1602
DB 1021 TCAACACTGGCCAGGCGAGTGAAGAAAGCTTGGGGAGAGAGCCAGCTGTATGGGGAGCCG 1080
QY 1603 TTCCAGCATGTCTTTCTACTGAGTGCAGAGAGCTGGCCAGTCCAGAGTGTGAGTCTC 1662
DB 1081 TTCCAGCATGTCTTTCTACTGAGTGCAGAGAGCTGGCCAGTCCAGAGTGTGAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCGGGCTCCCATTTAGACAGATCTGTCT 1722
DB 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCGGGCTCCCATTTAGACAGATCTGTCT 1200
QY 1723 AGGCGAGAGCGGCTGCTTTCATCTCGATGTGTAGATGAGCCAGATGGGCTTTCAG 1782
DB 1201 AGGCGAGAGCGGCTGCTTTCATCTCGATGTGTAGATGAGCCAGATGGGCTTTCAG 1260

QY 1783 GAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCAGCCACAGCCGGGAGATGACATGCTG 1842
DB 1261 GAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCAGCCACAGCCGGGAGATGACATGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAACTATACTTCCAGAGCATCTCTCTGATACAGGCTCGAGCC 1902
DB 1321 GGCAGTTTGTGGGAAAACTATACTTCCAGAGCATCTCTCTGATACAGGCTCGAGCC 1380
QY 1903 ACAGCTCTGCAAGACTCATCTCTTTTGGAGCAGGCAAGTGGGTAGAGTCTGGGG 1962
DB 1381 ACAGCTCTGCAAGACTCATCTCTTTTGGAGCAGGCAAGTGGGTAGAGTCTGGGG 1440
QY 1963 TTCTGAGTCCAGAGAAAGAAATTTCTACAGATTTTCAAGATGAAGCAAGCA 2022
DB 1441 TTCTGAGTCCAGAGAAAGAAATTTCTACAGATTTTCAAGATGAAGCAAGCA 1500
QY 2023 ATTAGAGCTTTAGTTGTGTCATTAATCAAAAGAGCTGTGGCCCTGTGTCTTGTGCC 2082
DB 1501 ATTAGAGCTTTAGTTGTGTCATTAATCAAAAGAGCTGTGGCCCTGTGTCTTGTGCC 1560
QY 2083 TGGGTGTCTGGCTGGCTTGCACTTGTCTGATGCAAGATGAAGCGGAAAGAAAACTC 2142
DB 1561 TGGGTGTCTGGCTGGCTTGCACTTGTCTGATGCAAGATGAAGCGGAAAGAAAACTC 1620
QY 2143 ACACTGACTTCCAGAGCCACCAACCCCTGTGTCTAATTACCTTGGCCAGGCTCTCCA 2202
DB 1621 ACACTGACTTCCAGAGCCACCAACCCCTGTGTCTAATTACCTTGGCCAGGCTCTCCA 1680
QY 2203 GCTCAGCATTTGGGAGCCCAAGCTCAGAGACCTTGTCTGTGCTGTGAGGCAATGG 2262
DB 1681 GCTCAGCATTTGGGAGCCCAAGCTCAGAGACCTTGTCTGTGCTGTGAGGCAATGG 1740
QY 2263 CAAAAAAGACCTTTTCACTGTCAGATGACCTCAGAGAGATGGGTAGATGGGGCCATC 2322
DB 1741 CAAAAAAGACCTTTTCACTGTCAGATGACCTCAGAGAGATGGGTAGATGGGGCCATC 1800
QY 2323 ATCTCAACCTCTTGAAGATGGGTATTTCTTCAAGAGACCCCAACCTCTGAGCTACAG 2382
DB 1801 ATCTCAACCTCTTGAAGATGGGTATTTCTTCAAGAGACCCCAACCTCTGAGCTACAG 1860
QY 2383 TTCAATCACCTCTGTTTCCAGAGTCTTGTGAGCAATGTCTATGTCTTGGAGATGAG 2442
DB 1861 TTCAATCACCTCTGTTTCCAGAGTCTTGTGAGCAATGTCTATGTCTTGGAGATGAG 1920
QY 2443 AAGGGAGAGTAAACATTTAATTGCATATGATTTGAAAAAGCGTAAAGCATAT 2502
DB 1921 AAGGGAGAGTAAACATTTAATTGCATATGATTTGAAAAAGCGTAAAGCATAT 1980
QY 2503 GGAATACATGGCCTGTTTGGGGGATCAACCAACGTTTCTAATTGGGCGTTAAAGTAT 2562
DB 1981 GGAATACATGGCCTGTTTGGGGGATCAACCAACGTTTCTAATTGGGCGTTAAAGTAT 2040
QY 2563 GAGGGGAGAGAGATGAGAAACATCTTTCATGCGGCTGTCTCAGGGAGAAACCTG 2622
DB 2041 GAGGGGAGAGAGATGAGAAACATCTTTCATGCGGCTGTCTCAGGGAGAAACCTG 2100
QY 2623 ATGCAATGGGTCCTGTCCTGCACTGCTGTGAGACCACTCTCTGAGTCCCTCAC 2682
DB 2101 ATGCAATGGGTCCTGTCCTGCACTGCTGTGAGACCACTCTCTGAGTCCCTCAC 2160
QY 2683 TGCTTGAACAGACTGAGAAACAAAGCTTCTGACCAAGTGAATGGCCATTTGAGAA 2742
DB 2161 TGCTTGAACAGACTGAGAAACAAAGCTTCTGACCAAGTGAATGGCCATTTGAGAA 2220
QY 2743 ATGGGAGTGTGTAGAAACAGACATGAGACTTATGTGACATTTCTGACATTAATTC 2802
DB 2221 ATGGGAGTGTGTAGAAACAGACATGAGACTTATGTGACATTTCTGACATTAATTC 2280
QY 2803 AGCCGCCACGTGAAGAAAGCTTCACTGATTTGAGGGGAGGAGACAGATCAATGAGAG 2862
DB 2281 AGCCGCCACGTGAAGAAAGCTTCACTGATTTGAGGGGAGGAGACAGATCAATGAGAG 2340
QY 2863 CCAACCATGTACTCCGTTCAGGTGGGTCCAGTCAAGATGCTAATTTGGCAGATTTCTC 2922

Db 2341 CCACACATGATATCTGTTTCAAGTGGTCCAGTCAAGATCCATTTGGACGATTTCT 2400
 Qy 2923 TTCTCGCTCTCAAGGTCAACAGAAACTGTAGAGAGCTGAGCTTAAGTGAACCTGCTG 2982
 Db 2401 TTCTCGCTCTCAAGGTCAACAGAAACTGTAGAGAGCTGAGCTTAAGTGAACCTGCTG 2460
 Qy 2983 AGCACTCTGCAATGAGAGATCTTTTGAACCTGTAGAGAGCTGCTGCTGCTGAG 3042
 Db 2461 AGCACTCTGCAATGAGAGATCTTTTGAACCTGTAGAGAGCTGCTGCTGCTGAG 2520
 Qy 3043 ACCCTGAGTGGCTGAGCTGAGCTCAAGTGAAGGCTTGAAGGCTTGGCTTGGG 3102
 Db 2521 ACCCTGAGTGGCTGAGCTGAGCTCAAGTGAAGGCTTGAAGGCTTGGCTTGGG 2580
 Qy 3103 CTGAGAGCCAAACAGACCTTGAACCTGAGCTGAGCTTGAAGTGTCTCAAGATGCT 3162
 Db 2581 CTGAGAGCCAAACAGACCTTGAACCTGAGCTGAGCTTGAAGTGTCTCAAGATGCT 2640
 Qy 3163 GAGAGCCAAACAGCTTTGAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAG 3222
 Db 2641 GAGAGCCAAACAGCTTTGAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAG 2700
 Qy 3223 CTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3282
 Db 2701 CTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2760
 Qy 3283 AGCCCAAGCTGAGAGAGCTGAGCTGAGAGAGCTGAGAGAGCTGAGAGCTGAG 3342
 Db 2761 AGCCCAAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 2820
 Qy 3343 CTGCTCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 3402
 Db 2821 CTGCTCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 2869
 Qy 3403 ACAACTCTGAGTATGATGATGAGAGAGAGCTGAGAGAGCTGAGAGAGAGCTGAG 3462
 Db 2870 ----- 2869
 Qy 3463 CTGCTCATCTTCAAG 3522
 Db 2870 -----GAG 2910
 Qy 3523 GAG 3582
 Db 2911 GAG 2970
 Qy 3583 GCTTCCATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3642
 Db 2971 GCTTCCATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3030
 Qy 3643 GCTGAGATGAG 3702
 Db 3031 GCTGAGATGAG 3090
 Qy 3703 TCTTCCGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3762
 Db 3091 TCTTCCGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3150
 Qy 3763 GAGCCCAAG 3822
 Db 3151 GAGCCCAAG 3210
 Qy 3823 CACTTCCCTGAG 3882
 Db 3211 TGGTTCCTCCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3270
 Qy 3883 GAG 3939
 Db 3271 GAG 3330
 Qy 3940 AACCCAG 3999

Db 3331 CAGACACATGAAACAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3390
 Qy 4000 GCTGTGAG 4059
 Db 3391 GAG 3450
 Qy 4060 TCCCTGTTCCAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4119
 Db 3451 TCCCTGTTCCAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3510
 Qy 4120 GTGAGAGCTGATACATAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4179
 Db 3511 GTGAGAGCTTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3570
 Qy 4180 AAAATGATTCATAG 4239
 Db 3571 CGATGAG 3630
 Qy 4240 CGATGAG 4299
 Db 3631 CAGCCCAAG 3690
 Qy 4300 CGAG 4360
 Db 3691 ACAAG 3704
 RESULT 11
 ID AAA78392 standard; cDNA; 2657 BP.
 AC AAA78392;
 DT 20-NOV-2000 (first entry)
 DE Human secreted protein gene 12 SEQ ID NO:22.
 KW Human; secreted protein; cytosolic; antianemic; antidiabetic;
 KW antiinflammatory; ophthalmological; antirheumatic; antiallergic;
 KW antiparasitic; antiparasitic; antiparasitic; antiparasitic;
 KW immunoprotective; antimicrobial; antiparasitic; antiparasitic;
 KW immune system disorder; angiogenesis; hyperproliferative disorder;
 KW cardiovascular disorder; apoptosis; neurological disease;
 KW infectious disease; wound healing; ss.
 OS Homo sapiens.
 PN MO200035937-A1.
 PD 22-JUN-2000.
 PF 16-DEC-1999; 99MO-US29950.
 PR 17-DEC-1998; 98US-0112809.
 PR 18-DEC-1998; 98US-0113006.
 PA (HUMA-) HUMAN GENOME SCT INC.
 PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
 PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
 DR P-PSDB; AAB24448.
 DR WPI; 2000-431566/37.
 XX Forty seven human nucleic acids encoding secreted proteins, useful in
 XX the treatment, prevention and diagnosis of cancers, disorders of the
 XX immune system, angiogenesis disorders, neurological diseases and
 XX hyperproliferative disorders -
 PS Claim 1; Page 449-450; 562bp; English.
 CC The polynucleotide sequence given in AAA78381 to AAA78432 encode the
 CC human secreted proteins given in AAB24437 to AAB24604. Human secreted

Db 1867 GCTGTTACCTTGAGCAGTCTCGGAGCCGGAAGTCAAGTGACTTACCAAGCCCT 1926
QY 4734 GAAGAGACCCATCTCACCCTCATTTAGAACTCTGGAGAAAGGAGAGCAAAAAGGAGCT 4793
Db 1927 GAAGAGACCCATCTCACCCTCATTTAGAACTCTGGAGAAAGGAGAGCAAAAAGGAGCT 1985
QY 4794 CCTGCACCTCAGACAGTGAAGTATCAACACGACCCCTTGAACCTTGAAGTCTTGCTTTGG 4853
Db 1986 CCGGCACCTCAGACAGTGAAGTATCAACACGACCCCTTGAACCTTGAAGTCTTGCTTTGG 2045
QY 4854 CTGACCCCTCTTGGGCTTCAGTTTCTTCTGCAAAACAAGTGGCATCTGGTTGGCT 4913
Db 2046 CTGACCCCTCTTGGGCTTCAGTTTCTTCTGCAAAACAAGTGGCATCTGGTTGGCT 2105
QY 4914 TCCAGACCTAAAGTATGAACTTTGATGATGCTTGGGACCTTATGTGTCATGCC 4973
Db 2106 TCCAGACCTAAAGTATGAACTTTGATGATGCTTGGGACCTTATGTGTCATGCC 2165
QY 4974 AGGATGCCACAGAGGGGCCCCAGTCCAGGTGCTTAACAGCATCTCAGGGAATGTCATC 5033
Db 2166 AGGATGCCACAGAGGGGCCCCAGTCCAGGTGCTTAACAGCATCTCAGGGAATGTCATC 2225
QY 5034 TGGAGCTGGCAAGACCCCTGCAAGCCTCATAGGCTCATCTGCTGGCCACAGCAACCA 5093
Db 2226 TGGAGCTGGCAAGACCCCTGCAAGCCTCATAGGCTCATCTGCTGGCCACAGCAACCA 2285
QY 5094 GCTTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAAATAGAGAGGACATGGAACCAT 5153
Db 2286 AGCTTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAAATAGAGAGGACATGGAACCAT 2345
QY 5154 TTGCTCTGCTGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGGTTCAAGCTGGAGGCCACG 5213
Db 2346 TTGCTCTGCTGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGGTTCAAGCTGGAGGCCACG 2405
QY 5214 TGGATCTTGCTGCTTTGATCAGAGAAATCTAACAAGACCAAGCAAGAGTAAAGTGAAG 5273
Db 2406 TGGATCTTGCTGCTTTGATCAGAGAAATCTAACAAGACCAAGCAAGAGTAAAGTGAAG 2465
QY 5274 GAAGTTATTCAGAAATTAAGAGATATCAGAGCTCTTTAGAAATTTGCTAGCAGGCTT 5333
Db 2466 GAAGTTATTCAGAAATTAAGAGATATCAGAGCTCTTTAGAAATTTGCTAGCAGGCTT 2525
QY 5334 TCCAGTTTATTCAGAAAACCCCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 5393
Db 2526 TCCAGTTTATTCAGAAAACCCCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2585
QY 5394 AAAAAATACAAAAAAGAAAAATGAATAATTAAGAAATTAAGAAATTAAGAAATTAAG 5444
Db 2586 AAAAAATACAAAAAAGAAAAATGAATAATTAAGAAATTAAGAAATTAAGAAATTAAG 2636

RESULT 12
AAL01744 standard, cDNA, 578 BP.
ID AAL01744;
AC AAL01744;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 1745.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN MO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01339.
XX

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0218963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226811.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0228343.
PR 01-SEP-2000; 2000US-0228344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-465570/50.
XX P-PSDB; AAM95774.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition.

XX	PS	Claim 1; SEQ ID NO 1745; 1297bp + Sequence Listing; English.
XX	CC	The present invention provides the protein and coding sequences of a
XX	CC	number of human reproductive system related antigens. These can be used
XX	CC	in the prevention and treatment of reproductive system disorders,
XX	CC	including cancer. The present sequence is a coding sequence of the
XX	CC	invention.
XX	XX	Sequence 578 BP, 130 A, 151 C, 166 G, 126 T, 5 other;
XX	XX	Query Match 8.9%; Score 484.4; DB 22; Length 578;
XX	XX	Best Local Similarity 99.4%; Pred. No. 5e-110;
XX	XX	Matches 496; Conservativity 0; Mismatches 2; Indels 1; Gaps 1;
QY	Db	1375 CAAAGACCTCACCCCAAGAGCCAAAGATCCCTGTGTCAAAGGAAGCTGGCTGATTATGTG 1434
QY	Db	2 CAAAGACCTCACCCCAAGAGCCAAAGATCCCTGTGTCAAAGGAAGCTGGCTGATTATGTG 61
QY	Db	1435 GAGAGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTGGCCAGGCTTGATATCC 1494
QY	Db	62 GAGAGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTGGCCAGGCTTGATATCC 122
QY	Db	1495 CAAAGACCTCGCATATGTCATATCTGCAAGG3G3GCTGCGAATTG3GAAGTCAACACTG3CC 1554
QY	Db	122 CAAAGACCTCGCATATGTCATATCTGCAAGG3G3GCTGCGAATTG3GAAGTCAACACTG3CC 181
QY	Db	1555 AGGAGAGTGAAGGAAGGCTGGGGGAGAGGCGAGCTGTATGGGAGCCGCTTCCAGCATGTC 1614
QY	Db	182 AGGAGAGTGAAGGAAGGCTGGGGGAGAGGCGAGCTGTATGGGAGCCGCTTCCAGCATGTC 241
QY	Db	1615 TTCTACTTCAGCTGCAAGAGAGCTGGCCAGTCCAAAGTGTGTAGTCTCGCTAGACTCATC 1674
QY	Db	242 TTCTACTTCAGCTGCAAGAGAGCTGGCCAGTCCAAAGTGTGTAGTCTCGCTAGACTCATC 301
QY	Db	1675 GGAAGAAATGGAGCAGCCATCTCCGGCTCCATTAGACAGATCCTGTCTATAGCCAGAGCCGG 1734
QY	Db	302 GGAAGAAATGGAGCAGCCATCTCCGGCTCCATTAGACAGATCCTGTCTATAGCCAGAGCCGG 361
QY	Db	1735 CTGCTCTTCATCCCTCGATGCTGTATGATGAGACCCAGATG3GCTTTGACAGAGCCGAGTTCT 1794
QY	Db	362 CTGCTCTTCATCCCTCGATGCTGTATGATGAGACCCAGATG3GCTTTGACAGAGCCGAGTTCT 421
QY	Db	1795 GAGCTCTGTGTCGACTGAGAGCCAGCCACAG-CCGGCGGATGCACTGCTGG3CAGTTTGTCT 1853
QY	Db	422 GAGCTCTGTGTCGACTGAGAGCCAGCCACAGCCGGCGGATGCACTGCTGG3CAGTTTGTCT 481
QY	Db	1854 GGGGAAAACATACTCTCC 1872
QY	Db	482 GGGGAAAACATACTCTCC 500
XX	XX	RESULT 13
XX	XX	ABL97037
XX	XX	ABL97037 standard; cDNA; 578 BP.
XX	XX	ABL97037;
XX	XX	21-JUN-2002 (first entry)
DE	DE	Human testicular antigen encoding cDNA SEQ ID NO: 705.
KM	KM	Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KM	KM	reproductive system disorder; urinary system disorder; gene therapy;
KM	KM	cardiovascular disorder; respiratory disorder; neurological disorder;
KM	KM	gastrointestinal disease; infection; cytoskeletal; gene; ss.
OS	OS	Homo sapiens.
XX	XX	WO200155317-A2.
XX	XX	02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01329.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0244675.
 PR 08-NOV-2000; 2000US-0244676.
 PR 08-NOV-2000; 2000US-0244677.
 PR 08-NOV-2000; 2000US-0244678.
 PR 08-NOV-2000; 2000US-024523.
 PR 08-NOV-2000; 2000US-024524.
 PR 08-NOV-2000; 2000US-024525.
 PR 08-NOV-2000; 2000US-024526.
 PR 08-NOV-2000; 2000US-024527.
 PR 08-NOV-2000; 2000US-024528.
 PR 08-NOV-2000; 2000US-024532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249247.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI, 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides.

PT useful for preventing, diagnosing and/or treating testicular cancer -
 XX
 PS Claim 1; SEQ ID NO 705; 766pp; English.

CC The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a cDNA of the
 CC invention.

XX
 SO Sequence 578 BP; 130 A; 151 C; 166 G; 126 T; 5 other;

Query Match 8.9%; Score 484.4; DB 23; Length 578;
 Best Local Similarity 99.4%; Pred. No. 5e-110;
 Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 CAAAGACTCCACCCCAAGAGCCAGATCCCTGGTCAAGAGAGAGCTGGCTGATTATGTG 1434
 DB 2 CAAAGACTCCACCCCAAGAGCCAGATCCCTGGTCAAGAGAGAGCTGGCTGATTATGTG 61

QY 1435 GAGGAGATCAGAGGACATTTATTTAGATCAGAGATTTATTTGGCCAGGCTGGATAC 1494
 DB 62 GAGGAGATCAGAGGACATTTATTTAGATCAGAGATTTATTTGGCCAGGCTGGATAC 121

QY 1495 CAAGAACCTCGCATAGTCATATCTCAGGGGGCTGCTGGAATTGGGAGTCAACTGAGCC 1554
 DB 122 CAAGAACCTCGCATAGTCATATCTCAGGGGGCTGCTGGAATTGGGAGTCAACTGAGCC 181

QY 1555 AGGCAAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCCAGCATATC 1614
 DB 182 AGGCAAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCCAGCATATC 241

QY 1515 TTCTACTTCACTGCAAGAGAGCTGGCCAGTCCAAAGGAGTGAAGTCTGCTGAGTATC 1674
 DB 242 TTCTACTTCACTGCAAGAGAGCTGGCCAGTCCAAAGGAGTGAAGTCTGCTGAGTATC 301

QY 1675 GGAAGAAGTGGAGACAGCCACTCCGGCTCCATTAGACAGATCTGTCTAGGCCAAGCGG 1734
 DB 302 GGAAGAAGTGGAGACAGCCACTCCGGCTCCATTAGACAGATCTGTCTAGGCCAAGCGG 361

QY 1735 CTGCTCTTCATCTTCGATGTGTAGATAGCCAGATGGGTCTTGACAGAGCCAGTTCT 1794
 DB 362 CTGCTCTTCATCTTCGATGTGTAGATAGCCAGATGGGTCTTGACAGAGCCAGTTCT 421

QY 1795 GAGCTCTGTGCACTGGAGCCAGCCAG--CCGGCGGATGACTGCTGGGCAATTGCT 1853
 DB 422 GAGCTCTGTGCACTGGAGCCAGCCAGCCCGGCGGATGACTGCTGGGCAATTGCT 481

QY 1854 GGGGAAAATTAATCTTCC 1872
 DB 482 GGGGAAAATTAATCTTCC 500

RESULT 14
 ABV97890/c
 ID ABV97890 standard; cDNA; 487 BP.

XX AC ABV97890;
 XX 14-JAN-2003 (first entry)

DE Human pancreatic cancer expressed cDNA SEQ ID NO 3298.
 XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
 KW cytostatic; tumour; gene; ss.

XX Homo sapiens.
 XX OS
 XX PN WO200260317-A2.
 XX PD 08-AUG-2002.

XX 30-JAN-2002; 2002WO-US02781.
 XX 30-JAN-2001; 2001US-265305P.
 XX 31-JAN-2001; 2001US-265682P.
 XX 09-FEB-2001; 2001US-267568P.
 XX 21-MAR-2001; 2001US-278651P.
 XX 28-APR-2001; 2001US-287112P.
 XX 16-MAY-2001; 2001US-305484P.
 XX 12-JUL-2001; 2001US-305484P.
 XX 20-AUG-2001; 2001US-313999P.
 XX 27-NOV-2001; 2001US-333626P.
 XX (CORI-) CORIYA CORP.
 XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
 XX WPI; 2002-627435/67.

XX New isolated polynucleotide and pancreatic tumor polypeptides, useful
 XX for diagnosing, preventing and/or treating cancer, particularly
 XX pancreatic cancer

XX Claim 1; SEQ ID NO 3298; 300pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising: (a)
 XX any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
 XX (b) complements of (a); (c) sequences consisting of at least 20
 XX contiguous residues of (a); (d) sequences that hybridize to (a), under
 XX moderately stringent conditions; (e) sequences having at least 75% or 90%
 XX identity to (a); or (f) degenerate variants of (a). Polypeptides
 XX (ABP6596-ABP6837) encoded by (1) and oligonucleotide can be used to
 XX detect cancer in a patient and compositions comprising polypeptides,
 XX polynucleotides, antibodies, fusion proteins, T cell populations and
 XX antigen presenting cells expressing the polypeptide are useful in
 XX treating pancreatic cancer and stimulating an immune response. The
 XX polynucleotides can be used as probes or primers for nucleic acid
 XX hybridisation, in the design and preparation of ribozyme molecules for
 XX inhibiting expression of the tumour polypeptides and proteins in the
 XX tumour cells, in vaccines and for gene therapy.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 487 BP; 114 A; 130 C; 133 G; 109 T; 1 other;

Query Match 8.9%; Score 482.8; DB 24; Length 487;
 Best Local Similarity 99.4%; Pred. No. 1.1e-109;
 Matches 484; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4744 CATCTCACTCATTAATGAACTGTGGAGAGGAGCAAAAAGGAGCTCTGCCACTC 4803
 DB 487 CATCTCACTCATTAATGAACTGTGGAGAGGAGCAAAAAGGAGCTCTGCCACTC 428

QY 4804 AGCAGCTGAAGTATCAACACAGCCCTTGAGCTTGTAGTCTTGCTTGGTGAACCTTC 4863
 DB 427 AGCAGCTGAAGTATCAACACAGCCCTTGAGCTTGTAGTCTTGCTTGGTGAACCTTC 368

QY 4864 TTGAGGTCTCAGTTCTTCTCTGCAAAAGAGTGCATCTGTTGGCTTCCAGACTA 4923
 DB 367 TTGAGGTCTCAGTTCTTCTCTGCAAAAGAGTGCATCTGTTGGCTTCCAGACTA 308

QY 4924 AAGTAATGAACCTTGAATGATGCTTTGCTGGGCAATTATGTTCATGCCAGGATGCCA 4983
 DB 307 AAGTAATGAACCTTGAATGATGCTTTGCTGGGCAATTATGTTCATGCCAGGATGCCA 248

QY 4984 CAGGGGGCCCCAGTCCAGTGGCCCTTAACAGATCTCAGGGAATGCCATCTGGAGCTGGC 5043
 DB 247 CAGGGGGCCCCAGTCCAGTGGCCCTTAACAGATCTCAGGGAATGCCATCTGGAGCTGGC 188

QY 5044 AAGAGCCCTGAGAGCTCATAGAGCTCATCTGTGGCCACAGAGCCAAAGCTTAGAGCC 5103
 DB 187 AAGAGCCCTGAGAGCTCATAGAGCTCATCTGTGGCCACAGAGCCAAAGCTTAGAGCC 128

QY 5104 CTCGGATCCCATCCAGCGCGAAGAGAAATAGAGGAGACATGAAACCATTTGCTCTGG 5163
 DB 127 CTCGGATCCCATCCAGCGCGAAGAGAAATAGAGGAGACATGAAACCATTTGCTCTGG 68
 QY 5164 CTGTGTACAGGGGTGAGCCCAAAATTTGGGTTCAAGGTGGAGGAGCCACGTGATTCTTG 5223
 DB 67 CTGTGTACAGGGGTGAGCCCAAAATTTGGGTTCAAGGTGGAGGAGCCACGTGATTCTTG 8
 QY 5224 GCTTTGT 5230
 DB 7 GCTTTGT 1
 RESULT 15
 ABV97891/c
 ID ABV97891 standard; cDNA; 416 BP.
 AC ABV97891;
 XX
 XX 14-JAN-2003 (first entry)
 DE Human pancreatic cancer expressed cDNA SEQ ID NO 3299.
 XX
 XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
 KM Cytostatic; tumour; gene; ss.
 XX
 XX Homo sapiens.
 PN WO200260317-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 30-JAN-2002; 2002WO-US02781.
 XX
 PR 30-JAN-2001; 2001US-265305P.
 PR 31-JAN-2001; 2001US-265682P.
 PR 09-FEB-2001; 2001US-267568P.
 PR 21-MAR-2001; 2001US-278651P.
 PR 28-APR-2001; 2001US-287112P.
 PR 16-MAY-2001; 2001US-291631P.
 PR 12-JUL-2001; 2001US-305484P.
 PR 20-AUG-2001; 2001US-313999P.
 PR 27-NOV-2001; 2001US-333626P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
 XX
 XX WPI; 2002-627435/67.
 DR
 XX
 PT New isolated polynucleotide and pancreatic tumor polypeptides, useful
 PT for diagnosing, preventing and/or treating cancer, particularly
 PT pancreatic cancer -
 PS
 PS Claim 1; SEQ ID NO 3299; 300bp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated polynucleotide (I) comprising: (a)
 CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
 CC (b) complements of (a); (c) sequences consisting of at least 20
 CC contiguous residues of (a); (d) sequences that hybridize to (a), under
 CC moderately stringent conditions; (e) sequences having at least 75% or 90%
 CC identity to (a); or (f) degenerate variants of (a). Polypeptides
 CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
 CC detect cancer in a patient and compositions comprising polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations and
 CC antigen presenting cells expressing the polypeptide are useful in
 CC treating pancreatic cancer and stimulating an immune response. The
 CC polynucleotides can be used as probes or primers for nucleic acid
 CC hybridisation, in the design and preparation of ribozyme molecules for
 CC inhibiting expression of the tumour polypeptides and proteins in the
 CC tumour cells, in vaccines and for gene therapy.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 416 BP; 101 A; 113 C; 114 G; 88 T; 0 other;
 Query Match 7.6%; Score 414.4; DB 24; Length 416;
 Best Local Similarity 99.8%; Pred. No. 1.1e-92;
 Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4815 TATCAACACGAGCCCTGACCCCTTGAGTCCTTGAGCTTGAGCCCTTTGGGTCCTCA 4874
 DB 416 TATCAACACGAGCCCTTGAGCCCTTGAGTCCTTGAGCTTGAGCCCTTTGGGTCCTCA 357
 QY 4875 GTTTCTTTCTCTGCAAAAGTTGCCATCTGTTGCCCTTCCAGCACTAAAGTAATGAA 4934
 DB 356 GTTTCTTTCTCTGCAAAAGTTGCCATCTGTTGCCCTTCCAGCACTAAAGTAATGAA 297
 QY 4935 CTTTGATGATGCTTTGCTGGGATATATGTCATGCCAGGATGCCAGAGGGGGCC 4994
 DB 296 CTTTGATGATGCTTTGCTGGGATATATGTCATGCCAGGATGCCAGAGGGGGCC 237
 QY 4995 AGTCCAGTGGCCTTAACAGCATCTCAGGGAATGTCATCTGAGCTGCAAGACCCCTGC 5054
 DB 236 AGTCCAGTGGCCTTAACAGCATCTCAGGGAATGTCATCTGAGCTGCAAGACCCCTGC 177
 QY 5055 AGACCTCATAGAGCCTCATCTGTGGCCACAGAGCCAAAGCCTTCCGATCCC 5114
 DB 176 AGACCTCATAGAGCCTCATCTGTGGCCACAGAGCCAAAGCCTTCCGATCCC 117
 QY 5115 ATCCAGGCGCAAGAGAAATAGAGGAGACATGAAACATTTGCTGCTGTGTACAG 5174
 DB 116 ATCCAGGCGCAAGAGAAATAGAGGAGACATGAAACATTTGCTGCTGTGTACAG 57
 QY 5175 GGTGAGCCCAAAATTTGGGTTCAAGGTGGAGGAGCCAGCTGATTTCTTGCTTTGT 5230
 DB 56 GGTGAGCCCAAAATTTGGGTTCAAGGTGGAGGAGCCAGCTGATTTCTTGCTTTGT 1

Search completed: January 29, 2004, 03:29:21
 Job time: 1357.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:01:52 ; Search time 10849.3 Seconds

(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-1

Perfect score: 5444

Sequence: 1 gccccagggcctgagagagc.....aggataagaagttactctac 5444

Scoring table: IDENTITY NUC
Gapd 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estma:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_estc3:*
12: gb_estc4:*
13: gb_estc5:*
14: gb_estc6:*
15: em_esthm:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	994	18.3	1201	9	AL541851
2	983.2	18.1	1201	13	BX341134
3	898.4	16.5	1201	13	BX341135
4	819.8	15.1	1055	13	BQ072636

5	814.4	15.0	1046	12	BM549709
6	787.4	14.5 <td>1110 <td>13</td> <td>BX457482</td> </td>	1110 <td>13</td> <td>BX457482</td>	13	BX457482
7	767	14.1 <td>768</td> <td>12</td> <td>B151887</td>	768	12	B151887
8	749.4	13.8	894	14	CD515485
9	727	13.4	791	10	BG682080
10	722.4	13.3	764	13	BO576345
11	671	12.3	686	14	CA308213
12	633.6	11.6	672	12	B1911134
13	633	11.5	728	12	B1908442
14	628.4	11.5	655	12	BQ003173
15	623.6	11.5	655	12	BM152462
16	615	11.3	615	12	BM791087
17	575	10.6	592	13	BU633947
18	566	10.4	566	10	BF059271
19	558.4	10.3	560	12	BM152839
20	544.4	10.0	554	12	BM193542
21	520.8	9.6	524	9	AM173055
22	520.4	9.6	523	12	BM153182
23	509	9.3	595	13	BO355077
24	502.6	9.2	514	9	AV682872
25	495.4	9.1	560	12	BM152807
26	494	9.1	494	9	AM593727
27	489.6	8.9	806	12	B1519523
28	485	8.9	494	12	BM151763
29	483.4	8.8	500	10	BE775300
30	476.4	8.8	633	12	B1909022
31	475	8.7	509	12	BG774227
32	474.6	8.7	521	9	A1819334
33	469.8	8.6	504	9	A1742596
34	469	8.6	469	9	A1825353
35	458	8.4	468	10	BE464469
36	456.8	8.4	511	13	BK474531
37	449	8.2	449	9	A1365110
38	445.4	8.2	461	9	AM291804
39	445.2	8.2	487	12	B1028123
40	441	8.1	464	9	AM291769
41	438	8.0	441	9	A1339112
42	434.4	8.0	466	12	BM193358
43	434.2	8.0	791	28	B2129303
44	424	7.8	425	9	A1439770
45	423.2	7.8	489	12	BM151640

ALIGNMENTS

RESULT 1
AL541851
LOCUS
DEFINITION
AL541851 Homo sapiens PLACENTA Homo sapiens cDNA clone CSDBE007YD22
ACCESSION
AL541851
VERSION
AL541851.2 GI:30546421
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12873320.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDBE007DB11QPLcluster=8396.r. Contact :
Peng Liang Email: fliang@lifetech.com URL :

QY		4426	TGGAGGCGCTTGTGTGAACCAAGAGATCTCAATGCCTGCAACTACTTCATCCCTCAGACC	4485
Db		997	TGTGGASGCCCTTGTAACAARAGANTCTCATGCTCGAACTACTTGATCCCTCAGACC	938
QY		4486	CGAATAGCCGTAACCTTCACTCTGTGATGCCCCGAGTTGCTGCATCTTTGTGGAACAAT	4545
Db		937	CGCATAGCCGTAACCTTCACTCTGTGATGCCCCGAGTTGCTGCATCTTTGTGGAACAAT	878
QY		4546	CGAGAGCAGCTGATAGCCCGAGTGCATCGGTGAGAGTTGTCTTGGACAAACTGCATGGA	4605
Db		877	CGAAGAGCAGCTGATAGCCCGAGTGCATCGGTGAGAGTTGTCTTGGACAAACTGCATGGA	818
QY		4606	CAGGTGCTGAGCCCAAGAGCACTACGAGAAGGTGTGTCTTGAGAACACAGAGCCCAGACCG	4665
Db		817	CAGGTGCTGAGCCCAAGAGCACTACGAGAAGGTGTGTCTTGAGAACACAGAGCCCAGACCG	758
QY		4666	ATGCGGAAGCTGTTCAAGCTTAGCCCAAGCTCTGGACCGGAAGTGCMAAGATGACTCTAC	4725
Db		757	ATGCGGAAGCTGTTCAAGCTTAGCCCAAGCTCTGGACCGGAAGTGCMAAGATGACTCTAC	698
QY		4726	CAAGCCTTGAAAGAGACCCATCTCACTCATTTATGAACTCTGGAGAAAGGACAGCAA	4785
Db		697	CAAGCCTTGAAAGAGACCCATCTCACTCATTTATGAACTCTGGAGAAAGGACAGCAA	638
QY		4786	AAGGAGCTCCCGCACATCAGAGCTGMAAGTATCAACACCAACCCTTGACCTTGAAGTCT	4845
Db		637	AAGGAGCTCCCGCACATCAGAGCTGMAAGTATCAACACCAACCCTTGAAGTCT	578
QY		4846	GGCATTGTGGCTGACCCCTTCTTGGGTCTCAGTTCTTCTCTGCAAAACAAGTTGCATCTG	4905
Db		577	GGCTTTGGCTGACCCCTTCTTGGGTCTCAGTTCTTCTCTGCAAAACAAGTTGCATCTG	518
QY		4906	GTTTGCCCTTCAGACACTPAAGTAAATGAACTTTGATGATGCTTGTCTGGCATTAATGTG	4965
Db		517	GTTTGCCCTTCAGACACTPAAGTAAATGAACTTTGATGATGCTTGTCTGGCATTAATGTG	458
QY		4966	TCCATGCAAGGGATAGCCAAGGGGGCCCACTCCAGAGTGGCTTAAACAGCATCTCAGGGAA	5025
Db		457	TCCATGCAAGGGATAGCCAAGGGGGCCCACTCCAGAGTGGCTTAAACAGCATCTCAGGGAA	398
QY		5026	TGTCATCTGTAGAGCTGGCAAGACCCCTTGACAGACTCATAGAGCCTCATCTGTGGCCACA	5085
Db		397	TGTCATCTGTAGAGCTGGCAAGACCCCTTGACAGACTCATAGAGCCTCATCTGTGGCCACA	338
QY		5086	GCAGCCCAAGCTTAAGACCCCTCCGATCCCATTCACAGCGCAAAAGAGAAATAGAGGACAT	5145
Db		337	GCAGCCCAAGCTTAAGACCCCTCCGATCCCATTCACAGCGCAAAAGAGAAATAGAGGACAT	278
QY		5146	GGAACCAATTTGCTCTGTGGCTGTGCACAGGGTGAAGCCCAAAATTTGGGGTTCAAGGTGGG	5205
Db		277	GGAACCAATTTGCTCTGTGGCTGTGCACAGGGTGAAGCCCAAAATTTGGGGTTCAAGGTGGG	218
QY		5206	AGGCGCAGGTGATTTCTTGGCTTTGTACAGAAATCTACAGAGCAACCAACAGATGA	5265
Db		217	AGGCGCAGGTGATTTCTTGGCTTTGTACAGAAATCTCTACAGAGCAACCAACAGATGA	158
QY		5266	AGTGAAGGAAGTTTATTCAGAAAAATAAGAGATATCAACACTCTTTTGAAGATTTGTCTA	5325
Db		157	AGTGAAGGAAGTTTATTCAGAAAAATAAGAGATATCAACACTCTTTTGAAGATTTGTCTA	98
QY		5326	GCAGGCTTCCAGTTTTTATTCAGAAAAACCCCTAATTAATTAATAATTTTACTTAATTT	5385
Db		97	GCAGGCTTCCAGTTTTTATTCAGAAAAACCCCTAATTAATTAATAATTTTACTTAATTT	38
QY		5386	AAGAAATTAATAATAATACAAAAGAA 5411	
Db		37	AAGAAATTAATAATAATACAAAAGAA 12	

DEFINITION	BX341135 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION	BX341135
VERSION	BX341135.1
KEYWORDS	EST.
ORGANISM	Homo sapiens (human)
SOURCE	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 1201)
TITLE	Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8396.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D012BB01QP1&cluster=8396.r. Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D012BB01QP1.
FEATURES	Location/Qualifiers
Source	1..1201
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0D012YC02"
	/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
	/cell_line="JURKAT"
	/clone_lib="Homo Sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT	297 a 329 c 312 g 217 t 46 others
ORIGIN	
Query Match	16.5%; Score 898.4; DB 13; Length 1201;
Best Local Similarity	97.4%; Pred. No. 6.2e-111;
Matches 937; Conservative	7; Mismatches 13; Indels 5; Gaps 3;
QY	773 CCCAAGGCCGGAAGGGGGAGGCGACCTCCCTCATTCCTCCCTACAGCCCAAGTAAACCC 832
DB	57 CCGGGGATCGAAGAGGGGGAGGCCACCTCTCTATTCCTTACAGCCCAAGTAAACCC 116
QY	833 ACCTGGGGGTCTCCAGGCAACCCACTCTCAACCGAGTGTAAATGCCCTGTATTCATGAT 892
DB	117 ACCTGGGGGTCTCCAGGCAACCCACTCTCAACCGAGTGTAAATGCCCTGTATTCATGAT 176
QY	893 TGCCTGGGGGTGTACCCAGGGCTCAGAGAGAAAGGTTTGAAGCAGCTGCTGACACAT 952
DB	177 TGCCTGGGGGTGTACCCAGGGCTCAGAGAGAAAGGTTTGAAGCAGCTGCTGACACAT 236
QY	953 CTGGAACCCCGTGGAGGAATCTCTGCTCACTCCCTTACCAAGCTCTTCAAGCTCC 1012
DB	237 CTGGAACCCCGTGGAGGAATCTCTGCTCACTCCCTTACCAAGCTCTTCAAGCTCC 296
QY	1013 CAGACATGATCTCCAGGCGAGGTACCCAGCGCCCCCACTCAAGAGCTGTGTG 1072
DB	297 CAGACATGATCTCCAGGCGAGGTACCCAGCGCCCCCACTCAAGAGCTGTGTG 356
QY	1073 GGAAGCTGGGGATCCCACTCAGCCGAGCTAGACCCAGAGAGAGAGAGGCTCTGGGA 1132
DB	357 GGAAGCTGGGGATCCCACTCAGCCGAGCTAGACCCAGAGAGAGAGAGGCTCTGGGA 416
QY	1133 CCGAATGGCTCTGGAAGAAACGTCAAGAAATTTATCTACACAAGAAATGAGAAAGAGA 1192
DB	417 CCGAATGGCTCTGGAAGAAACGTCAAGAAATTTATCTACACAAGAAATGAGAAAGAGA 476

```

QY 1193 GAGAGAAATCAGAGAAAGGAGGCCCCCATGAGGCGGTGTAGGAACGCCCAACAG 1252
DB 477 GAGAGAAATCAGAGAAAGGAGGCCCCCATGAGGCGGTGTAGGAACGCCCAACAG 536
QY 1253 CGACACCGAGCCCTACAGCCCCCAACCAACCAAGGAGGCTTCTGTAGAGAGAGGCTCT 1312
DB 537 CGACACCGAGCCCTACAGCCCCCAACCAACCAAGGAGGCTTCTGTAGAGAGAGGCTCT 596
QY 1313 GTTCCACATGAGCCCTGTAGAAAATGAGGATTTTAAACAAAATTACACAGCTCTACTTC 1372
DB 597 GTTCCACATGAGCCCTGTAGAAAATGAGGATTTTAAACAAAATTACACAGCTCTACTTC 656
QY 1373 TACAAAGACCTCACCCCAAGAGCCCAAGATCCCTGTGCAAGAGAGAGCTGCTGATTATG 1432
DB 657 TACAAAGACCTCACCCCAAGAGCCCAAGATCCCTGTGCAAGAGAGAGCTGCTGATTATG 716
QY 1433 TGGAGAGAGATGAGAGACATTTAATTGATGATCAGACATTTATTTGCCCCAGGCTGAGTA 1492
DB 717 TGGAGAGAGATGAGAGACATTTAATTGATGATCAGACATTTATTTGCCCCAGGCTGAGTA 776
QY 1493 CCCAAGAACCTCGCATAGTCTATCTGACAGGAGGCTGCTGGAATTGGAAATCAACACTGG 1552
DB 777 CCCAAGAACCTCGCATAGTCTATCTGACAGGAGGCTGCTGGAATTGGAAATCAACACTGG 836
QY 1553 CCAGGACGTTGAAGAGAGCCCTGGGAGAGAGGCGCAGCTGTATGGGGAACCGCTTCAGCATG 1612
DB 837 CCAGGACGTTGAAGAGAGCCCTGGGAGAGAGGCGCAGCTGTATGGGGAACCGCTTCAGCATG 896
QY 1613 TCTTCTAATTGAGCTGACAGAGAGCTGCCCCAGTCCAAAGGTGTGAGTCTTCTGAGACTCA 1672
DB 897 TCTTCTAATTGAGCTGACAGAGAGCTGCCCCAGTCCAAAGGTGTGAGTCTTCTGAGACTCA 955
QY 1673 TCGGAAAGATGAGAGAGCCACTCCGAGTCCCATTTAGACAGATCTGTCTAGGCGCAGAG 1732
DB 956 TCGGAAAGATGAGAGAGCCACTCCGAGTCCCATTTAGACAGATCTGTCTAGGCGCAGAG 1011
QY 1733 GG 1734
DB 1012 GS 1013

RESULT 4
LOCUS B0072636 1055 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6763126 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5756099
5', mRNA Sequence.
ACCESSION B0072636 GI:19901682
VERSION B0072636.1 GI:19901682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1055)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM12796 row: C column: 12
High quality sequence stop: 692.
Location/Qualifiers
1..1055
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db xref="taxon:9606"
/clone="IMAGE:5756099"
/tissue_type="Leukocyte"
/lab host="DH10B"
/clone_1ib="NIH MGC 118"
/notes="Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV
(deleted); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC Library."
BASE COUNT 220 a 349 c 282 g 202 t 2 others
ORIGIN

Query Match 15.1%; Score 819.8; DB 13; Length 1055;
Best Local Similarity 98.7%; Pred. No. 2.1e-100;
Matches 846; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1 GCCCAGGAGCTGAGAGGCTGAGAGAACTGGAGCAGAGCCGGGCTCCACTCT 60
DB 50 GCCCAGGAGCTGAGAGGCTGAGAGAACTGGAGCAGAGCCGGGCTCCACTCT 109
QY 61 GGGTCTGAAAGCCCATTCCTGCTCTGCGGCTCTCCACCCACCTTCTCAGCCTT 120
DB 110 GGGTCTGAAAGCCCATTCCTGCTCTGCGGCTCTCCACCCACCTTCTCAGCCTT 169
QY 121 GCAGCTCAAGGGTGTATCTCAGAGATCCAGAGCCCAAGAGAGAGAGAAATCTAGAGAA 180
DB 170 GCAGCTCAAGGGTGTATCTCAGAGATCCAGAGCCCAAGAGAGAGAGAAATCTAGAGAA 229
QY 181 CAGAAAGAGAGAGCTTGAGAGCCCAACCTCTCCCTCAACATCTCCCTCAACCTTCA 240
DB 230 CAGAAAGAGAGAGCTTGAGAGCCCAACCTCTCCCTCAACATCTCCCTCAACCTTCA 289
QY 241 CCTCCCTGCTGAGCCCTGAGAGCCCAACCTCTCCCTCAACATCTCCCTCAACCTTCA 300
DB 290 CCTCCCTGCTGAGCCCTGAGAGCCCAACCTCTCCCTCAACATCTCCCTCAACCTTCA 349
QY 301 GTCTTGACAGCCCTCTGAGGCTCTCCCTCCCTCTTCTCAACATCTCCCTCAACCTTCA 360
DB 350 GTCTTGACAGCCCTCTGAGGCTCTCCCTCCCTCTTCTCAACATCTCCCTCAACCTTCA 409
QY 361 CGAGCTCTATCTGAGAGTCCCTGAGGATTTATTAAGTGGTTCGGAATGTAATAAG 420
DB 410 CGAGCTCTATCTGAGAGTCCCTGAGGATTTATTAAGTGGTTCGGAATGTAATAAG 469
QY 421 GACGGTAAAGCCCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 480
DB 470 GACGGTAAAGCCCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 529
QY 481 CTGGGAACATCTCCCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 540
DB 530 CTGGGAACATCTCCCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 589
QY 541 GGCCTGCTGCTGTTACTTGAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 590 GGCCTGCTGCTGTTACTTGAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649
QY 601 CTGCTGCAATTAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAG 660
DB 650 CTGCTGCAATTAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAG 709
QY 661 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 710 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 769
QY 721 TGGAGACTAGCCCTCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
DB 770 TGGAGACTAGCCCTCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 829

```


http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0CAP008CG080P1&cluster=8396.r. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fullength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0CAP008CG080P1.
 Location/Qualifiers
 1. 1110
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP008YX15"
 /issue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (at) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

BASE COUNT 218 a 270 c 257 g 262 t 103 others
 ORIGIN

Query Match 14.5%; Score 787.4; DB 13; Length 1110;
 Best Local Similarity 88.3%; Pred. No. 4.6e-96;
 Matches 848; Conservative 58; Mismatches 48; Indels 6; Gaps 6;

1644 GTCCAGAGTGTAGTCTCGCTAGCTCATCGGAAAAGATGGAGCCACTCCGGCTCC 1703
 72 GACCCACGGGTCGGGTCCTCGCTAGCTCATCGGAAAAGATGGAGCCACTCCGGCTCC 131
 1704 CATTAGACAGATCTGTCTAGGCGCAGAGCGGCTGCTTTCATCTCGATGGTGTAGATGA 1763
 132 CATTAGACAGATCTGTCTAGGCGCAGAGCGGCTGCTTTCATCTCGATGGTGTAGATGA 191
 1764 GCCAGAGTGGTCTTTCAGAGAGCGGAGTTCTAGCTCTGTCTGACCTGAGCCAGCCACA 1823
 192 GCCAGAGTGGTCTTTCAGAGAGCGGAGTTCTAGCTCTGTCTGACCTGAGCCAGCCACA 251
 1824 GCGCGGAGTGTAGCTGTGGGCGATTTGCTGGGGAAAACATATCTTCCGAGGATCTT 1883
 252 GCGCGGAGTGTAGCTGTGGGCGATTTGCTGGGGAAAACATATCTTCCGAGGATCTT 311
 1884 CCGATACAGCGGCTCGACCAAGCTCTGCAAGACCTCATCTCTTGGAGAGGACAG 1943
 312 CCGATACAGCGGCTCGACCAAGCTCTGCAAGACCTCATCTCTTGGAGAGGACAG 371
 1944 TTGGGTAGAGGTCTTGGGATCTCTCTAGTCCAGCAGAGAGAAATTTCTACAGATATT 2003
 372 TTGGGTAGAGGTCTTGGGATCTCTCTAGTCCAGCAGAGAGAAATTTCTACAGATATT 431
 2004 CACAGATGAAGGCAAGCAATTAGAGCTTTAGGTTGTCAATCAAAACAAGGCTCTG 2063
 432 CACAGATGAAGGCAAGCAATTAGAGCTTTAGGTTGTCAATCAAAACAAGGCTCTG 491
 2064 GGGCCTGTGTCTTGTGCGCTGGGATCTCTGAGCTGAGCTTGGCTTGAATGAGAGAT 2123
 492 GGGCCTGTGTCTTGTGCGCTGGGATCTCTGAGCTTGGCTTGAATGAGAGAT 551
 2124 GAAGCGAAGGAAAACCTCACTGACTTCAAGACCAACCAACCTCTGTGTACATTA 2183
 552 GAAGCGAAGGAAAACCTCACTGACTTCAAGACCAACCAACCTCTGTGTACATTA 611
 2184 CTTTGGCCAGGCTCTCAAGCTCAGCAATTGGGACCCCACTCAGAGAA-CCTCTGCTTC 2242
 612 CTTTGGCCAGGCTCTCAAGCTCAGCAATTGGGACCCCACTCAGAGAA-CCTCTGCTTC 671
 2243 TGGCTGTGAGGAGCATCTGGCAAAA-AGACCTTTTTCAGTCCAGATGACTCAGAGAG 2301
 672 TGGCTGTGAGGAGCATCTGGCAAAA-AGACCTTTTTCAGTCCAGATGACTCAGAGAG 731
 2302 CATGGG-TTATGATGGGCAATCATCTCCACTTTTGAAGATGGGATTTCTTCAAGACA 2360
 732 CATGGGTTTATGATGGGCAATCATCTCCACTTTTGAAGATGGGATTTCTTCAAGACA 791

2361 CCCATCCCTCTGAGCTACAGCTTCACTCACCTCTGTTTCCAGAGT-CTTTGAGACA 2419
 DB 792 CCCATCCCTCTGAGCTACAGCTTCACTCACCTCTGTTTCCAGAGT-CTTTGAGACA 851
 2420 TGTCTATGTCTTGGAGATGAGAGGAGAGGATTAACATTTATTCATGATCATAGA-T 2478
 DB 852 TGTCTATGTCTTGGAGATGAGAGGAGAGGATTAACATTTATTCATGATCATAGA-T 911
 2479 TTGAAAAGAGCGCTAGA-AGCATATGAAATACATGCGCTGTTGGGCAATCAACACAG 2537
 DB 912 TTGAAAAGAGCGCGCTTTCATGAAATACATGCGCTGTTGGGCAATCAACACAG 971
 2538 TTTCTATGTGCGCTGTTAAGATGAGAGGAGAGAGAGATGAGAAATCTTCACTG 2597
 DB 972 TTGATATGCGGCGCTTTCATGAAATACATGCGCTGTTGGGCAATCAACACAG 1031

RESULT 7
 LOCUS B1518887 768 bp mRNA linear EST 29-AUG-2001
 DEFINITION B1518887 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211072 5',
 mRNA sequence.
 ACCESSION B1518887
 VERSION B1518887.1 GI:15343679
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 768)
 NIH-MGC http://mgi.nhl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL1529 row: n column: 01
 High quality sequence stop: 766.
 Location/Qualifiers

FEATURES
 source
 1. 768
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5211072"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."

BASE COUNT 176 a 213 c 225 g 154 t
 ORIGIN

Query Match 14.1%; Score 767; DB 12; Length 768;
 Best Local Similarity 100.0%; Pred. No. 2.7e-93;
 Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3119 CCTGACCGAGCTGACCTGAGCTTCAATGTGTCTACGAGATGTGAGACCAACACTTT 3178
 DB 1 CCTGACCGAGCTGACCTGAGCTTCAATGTGTGTCTACGAGATGTGAGACCAACACTTT 60

Accession	Source	Organism	Definition	Accession	Source	Organism	Definition
3179	CCNA	Homo sapiens	CCNA3 gene, complete cds	3179	CCNA	Homo sapiens	CCNA3 gene, complete cds
61	CCNA	Homo sapiens	CCNA3 gene, complete cds	61	CCNA	Homo sapiens	CCNA3 gene, complete cds
3239	CCNA	Homo sapiens	CCNA3 gene, complete cds	3239	CCNA	Homo sapiens	CCNA3 gene, complete cds
121	CCNA	Homo sapiens	CCNA3 gene, complete cds	121	CCNA	Homo sapiens	CCNA3 gene, complete cds
3299	CCNA	Homo sapiens	CCNA3 gene, complete cds	3299	CCNA	Homo sapiens	CCNA3 gene, complete cds
181	CCNA	Homo sapiens	CCNA3 gene, complete cds	181	CCNA	Homo sapiens	CCNA3 gene, complete cds
3359	CCNA	Homo sapiens	CCNA3 gene, complete cds	3359	CCNA	Homo sapiens	CCNA3 gene, complete cds
241	CCNA	Homo sapiens	CCNA3 gene, complete cds	241	CCNA	Homo sapiens	CCNA3 gene, complete cds
3419	CCNA	Homo sapiens	CCNA3 gene, complete cds	3419	CCNA	Homo sapiens	CCNA3 gene, complete cds
301	CCNA	Homo sapiens	CCNA3 gene, complete cds	301	CCNA	Homo sapiens	CCNA3 gene, complete cds
3479	CCNA	Homo sapiens	CCNA3 gene, complete cds	3479	CCNA	Homo sapiens	CCNA3 gene, complete cds
361	CCNA	Homo sapiens	CCNA3 gene, complete cds	361	CCNA	Homo sapiens	CCNA3 gene, complete cds
3539	CCNA	Homo sapiens	CCNA3 gene, complete cds	3539	CCNA	Homo sapiens	CCNA3 gene, complete cds
421	CCNA	Homo sapiens	CCNA3 gene, complete cds	421	CCNA	Homo sapiens	CCNA3 gene, complete cds
3599	CCNA	Homo sapiens	CCNA3 gene, complete cds	3599	CCNA	Homo sapiens	CCNA3 gene, complete cds
481	CCNA	Homo sapiens	CCNA3 gene, complete cds	481	CCNA	Homo sapiens	CCNA3 gene, complete cds
3659	CCNA	Homo sapiens	CCNA3 gene, complete cds	3659	CCNA	Homo sapiens	CCNA3 gene, complete cds
541	CCNA	Homo sapiens	CCNA3 gene, complete cds	541	CCNA	Homo sapiens	CCNA3 gene, complete cds
3719	CCNA	Homo sapiens	CCNA3 gene, complete cds	3719	CCNA	Homo sapiens	CCNA3 gene, complete cds
601	CCNA	Homo sapiens	CCNA3 gene, complete cds	601	CCNA	Homo sapiens	CCNA3 gene, complete cds
3779	CCNA	Homo sapiens	CCNA3 gene, complete cds	3779	CCNA	Homo sapiens	CCNA3 gene, complete cds
661	CCNA	Homo sapiens	CCNA3 gene, complete cds	661	CCNA	Homo sapiens	CCNA3 gene, complete cds
3839	CCNA	Homo sapiens	CCNA3 gene, complete cds	3839	CCNA	Homo sapiens	CCNA3 gene, complete cds
721	CCNA	Homo sapiens	CCNA3 gene, complete cds	721	CCNA	Homo sapiens	CCNA3 gene, complete cds

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
<http://image.jnl.gov>
plate: NDNA488 row: d column: 02
High quality sequence step: 606.
Location/Qualifiers

```

BASIC COUNT      194 a
ORIGIN            241 C      268 g      189 t      2 others

1. .854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30407209"
/cisue_type="White Matter"
/dev_stage="unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)
/clone_1ib="NIH_MGC_181"
/note="Vector: PCMV-SPORT6.1; Site_1: NotI; Site_2: E
(destroyed); Library is oligo-dT primed and direction
cloned (BCORV site is destroyed upon cloning) . Average
insert size 1.42 kb. Library was constructed by
(invitrogen). Note: this is a NIH_MGC Library."

```

Query Match	13.8%	Score	749.4	DB	14	Length	894
Best Local Similarity	95.7%	Pred. No.	66-91				
Matches	780	Conservative	0	Mismatches	33	Indels	2
						Gaps	1

QY	3202	TGCAGGCTACAGCGACTTCAGCTGCTCAGCTGTGGCTCAGCTGCACAGGAC	3267
Db	1	TGCAGGCTACAGCGACTTCAGCTGCTCAGCTGTGGCTCAGCTGCACAGGAC	60
QY	3262	CTGGCCCTCTGCTGCTTATATGCGACGCCCAAGCTGAGAGGACTAGAGCTGACAGAAACAAC	3322
Db	61	CTGGCCCTCTGCTGCTTATATGCGACGCCCAAGCTGAGAGGACTAGAGCTGACAGAAACAAC	120
QY	3322	CTGATAGACGTTGGCGTGCACATCTGCTGTGAGGGGCTCAGGCACTCTGCCTGCACAACTC	3381
Db	121	CTGATAGACGTTGGCGTGCACATCTGCTGTGAGGGGCTCAGGCACTCTGCCTGCACAACTC	180
QY	3382	ATTAGCCTGGGGGCTGAGCCAGACAACTCTGATGATGATGAGGACGAGAACTGAGGGCC	3442
Db	181	ATTAGCCTGGGGGCTGAGCCAGACAACTCTGATGATGATGAGGACGAGAACTGAGGGCC	240
QY	3442	CTGAGACGAGAGAAACCTCAGCTGCTCATCTTCAGCAGACGAGAACCAAGTGTATGACC	3501
Db	241	CTGAGACGAGAGAAACCTCAGCTGCTCATCTTCAGCAGACGAGAACCAAGTGTATGACC	300
QY	3502	CTTACTGAGGGGCTTGATACGGGAGAGATGATATATGACATCTCTCACTCAAGCGGCGAG	3561
Db	301	CTTACTGAGGGGCTTGATACGGGAGAGATGATATATGACATCTCTCACTCAAGCGGCGAG	360
QY	3562	AGACTCGGATCAGAGAGGGCGGCTTCCCATGTTTCTCAGGCTTAATCTCAAACTCCTGGAC	3621
Db	361	AGACTCGGATCAGAGAGGGCGGCTTCCCATGTTTCTCAGGCTTAATCTCAAACTCCTGGAC	420
QY	3622	GTCAGCAAGATCTTCCCAATTGCTGAGATTGTCAGAGGAAAGCTCCCAAGAGTGTATCG	3681
Db	421	GTCAGCAAGATCTTCCCAATTGCTGAGATTGTCAGAGGAAAGCTCCCAAGAGTGTATCG	480
QY	3682	GTCGGAATCTTTGGTCGTCGCTCTCCGCGCTCTCAAGGGGACCTGCATACGAAGCCTTTG	3741
Db	481	GTCGGAATCTTTGGTCGTCGCTCTCCGCGCTCTCAAGGGGACCTGCATACGAAGCCTTTG	540
QY	3742	GCGACTACGATGACTCTTCTGCGGGCCCAACGCGGGCTGTGGCTACTGAGAGTGTGACAA	3801
Db	541	GCGACTACGATGACTCTTCTGCGGGCCCAACGCGGGCTGTGGCTACTGAGAGTGTGACAA	600
QY	3802	GAAAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGACCAACACG	3861
Db	601	GAAAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGACCAACACG	660
QY	3862	GGTCTCTGCTTTGTGTATGAGAGAAAGCGGTGACCGTTGATGATTGAATTCGTGTGTGGGAC	3921

Db 661 GGTCTGCTTGTGTGATGAGAGAGGCTGACCGTTGAGATTGAATTCTGTGTGGAC 720

Qy 3922 CAGTT--CTGGGTGATGATCAACCAAGACAGCTGTGATGTGCGAGGCTCTGCTGG 3979

Db 721 CAGTTCTGGGGGTGATGATCAACCAAGACAGCTGTGATGTGCGAGGCTCTGCTGG 780

Qy 3980 ACATCAAGGCTGAGCGTGGAGCTGTGAAAGCTGTG 4014

Db 781 TGGNACATCAAGGCTTGAACCTGGAGAGCTGTG 815

RESULT 9
Bg682080 791 bp mRNA linear EST 01-MAY-2001
LOCUS 602630057F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754607 5',
DEFINITION mRNA sequence.
ACCESSION Bg682080 GI:13913477
VERSION Bg682080.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS 1 (bases 1 to 791)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: LLM10616 row: b. column: 16
High quality sequence stop: 770.
Location/Qualifiers
1. 791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754607"
/cissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1b="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 195 a 203 c 200 g 193 t

ORIGIN

Query Match 13.4%; Score 727; DB 10; Length 791;
Best Local Similarity 98.2%; Pred. No. 6.1e-88;
Matches 777; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

Qy 2045 AATCAACAAGAGCTCTGGGCGCTGTGTGCTGTGGCTGTGGCTGTGGCTGTGCA 2104

Db 3 AATCAACAAGAGCTCTGGGCGCTGTGTGCTGTGGCTGTGGCTGTGGCTGTGCA 61

Qy 2105 CTTCCTGTATGACGACGATGAGCGGAGAGAGAGAAATCTCACTGACTTCCAGACCA 2164

Db 62 CTTCCTGTATGACGACGATGAGCGGAGAGAGAGAAATCTCACTGACTTCCAGACCA 121

Qy 2165 CAACCTCTGTCTACATTAAGCTTGGCCAGGCTTCCAGAGTCAAGCATTTGGAGCCAGC 2224

Db 122 CAACCTCTGTCTACATTAAGCTTGGCCAGGCTTCCAGAGTCAAGCATTTGGAGCCAGC 181

Qy 2225 TCAGAGACCTGTCTGTGCTGTGAGGCGCATCTGGCAAAAAGACCTTTTCAGTC 2284

Db 182 TCAGAGACCTGTCTGTGCTGTGAGGCGCATCTGGCAAAAAGACCTTTTCAGTC 241

Qy 2285 CAGATGACCTCAAGAAAGATGGGTTGATGGGGCCATCTCTCCACTTTCTTGAAGATGG 2344

Db 242 CAGATGACCTCAAGAAAGATGGGTTGATGGGGCCATCTCTCCACTTTCTTGAAGATGG 301

Qy 2345 GTATCTTCAAGAGACCCCATCCCTGTGAGGTCAAGCTTCACTTCACTCTGTTCGAAG 2404

Db 302 GTATCTTCAAGAGACCCCATCCCTGTGAGGTCAAGCTTCACTTCACTCTGTTCGAAG 361

Qy 2405 AGTTCTTTGACAGCAATGCTCTATGCTTTGAGAGATGAGAGGGAGAGATTAACATTCTA 2464

Db 362 AGTTCTTTGACAGCAATGCTCTATGCTTTGAGAGATGAGAGGGAGAGATTAACATTCTA 421

Qy 2465 ATTGCATCATTAATTTGAAAAAGACGCTTGAAGCATATGGAATTAATGAGCTTTGGGG 2524

Db 422 ATTGCATCATTAATTTGAAAAAGACGCTTGAAGCATATGGAATTAATGAGCTTTGGGG 481

Qy 2525 CATCAACCAACAGTTTCCATTTGGGGCTGTTAAGTATGAGGGGGAGAGAGATGAGAGA 2584

Db 482 CATCAACCAACAGTTTCCATTTGGGGCTGTTAAGTATGAGGGGGAGAGAGATGAGAGA 541

Qy 2585 ACATCTTTTCACTGCCGCTGTCTCAGGGAGGAGACCTGATGAGTGGTCCGCTCCG 2644

Db 542 ACATCTTTTCACTGCCGCTGTCTCAGGGAGGAGACCTGATGAGTGGTCCGCTCCG 601

Qy 2645 ACCTGCTGTGACGACCACTTCTGAGAGTCCCTCCACTGCTTGTGACGAGTCCGAGACA 2704

Db 602 ACCTGCTGTGACGACCACTTCTGAGAGTCCCTCCACTGCTTGTGACGAGTCCGAGACA 661

Qy 2705 AAACGTTCTGTACACAGATGATGAGCCATTTGAGAAATAGGGCA-TGAGTGTAGAAACA 2763

Db 662 AAA-GTTCTGTACACAGATGATGAGCCATTTTGAAGAAATAGGGCA-TGAGTGTAGAAACA 720

Qy 2764 GACATGAGCTTTAGTGTGACCTTTTGCATT-AAATTACGCCCGCACGTGAAGAGCT 2822

Db 721 GACATGAGCTTTAGTGTGACCTTTTGCATTAAATTACGCCCGCACGTGAAGAGCT 780

Qy 2823 TCAGCTGATTG 2833

Db 781 TCAGCTGATTG 791

RESULT 10
Bg576345/c 764 bp mRNA linear EST 19-JUN-2002
LOCUS BQ576345
DEFINITION UI-H-E21-dbl-n-10-0-UI-s1 NCI CGAP Ch2 Homo sapiens cDNA clone
UI-H-E21-dbl-n-10-0-UI 3', mRNA sequence.
ACCESSION BQ576345 GI:21479662
VERSION BQ576345.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS 1 (bases 1 to 764)
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopedics
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-71, >Poly_A#Simple_repeat (matched complement) 72-100,
>AT-rich#low complexity 157-242, >MER41B#LTR/MER4-9group
Seq primer: M13 FORWARD

Db 540 GGAGCTGCATCATAGTCTTGAAAACCCAGCTTCTCCCTTGGAGTCTCTTGA 599

QY 4182 AATATCCATATAGCCCTGAGCTTCATTCGGTCAACCTGNGGNTGCTTACACCG 4241

Db 600 AATATCCATATAGCCCTGAGCTTCATTCGGTCAACCTGNGGNTGCTTACACCG 659

QY 4242 CGTC 4245

Db 660 CGTC 663

RESULT 13
Bi908442 728 bp mRNA linear EST 16-OCT-2001
LOCUS Bi908442
DEFINITION 603067393P1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216327 5',
mRNA sequence.
ACCESSION Bi908442
VERSION Bi908442.1 GI:16171396
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 728)
NIH-MGC <http://mgi.nci.nih.gov/>.
Natlional Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgabs-remail.nih.gov
COMMENT Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM1543 row: h column: 24
High quality sequence stop: 718.
Location/Qualifiers
1..728
source /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5216327"
/issue_type="leukocyte"
/lab_host="DH10B"
/clone_id="NIH_MGC_118"
/note="Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC Library."

BASE COUNT 200 a 208 c 200 g 120 t

ORIGIN

Query Match 11.6%; Score 633; DB 12; Length 728;
Best local similarity 98.3%; Pred. No. 2.4e-75;
Matches 713; Conservative 0; Mismatches 5; Indels 7; Gaps 7;

QY 885 CCATGAATTCGGCGGGGTCACCCAGGCTCAGAGAGAGGTTTGAACAGCTGCC 944

Db 6 CCATGAATTCGGCGGGGTCACCCAGGCTCAGAGAGAGGTTTGAACAGCTGCC 64

QY 945 TGACACATCTGGAAGCGCTGAGAGAAATCTGCGCTCACTCAACAGGCTTCC 1004

Db 65 TGACACATCTGGAAGCGCTGAGAGAAATCTGCGCTCACTCAACAGGCTTCC 124

QY 1005 AAGCTCCAGACCATGATCTCCAGGAGAGTCAACCAAGCCCAATCCACAGC 1064

Db 125 AAGTCCCCAGACCATGATCTCCAGGCCAGAGTCAACCAAGCCCCACATCCACAGC 184

QY 1065 AGTCTGGGAGAGCTGGGGATCCCCACCTCAGGCCCTAGACCCAGAGAGAGAGC 1124

Db 185 AGTCTGGGAGAGCTGGGGATCCCCACCTCAGGCCCTAGACCCAGAGAGAGAGC 244

QY 1125 TCCTGGAGACCAATGAGCTCTGATGAACGTCAGAAATTTATCAACAGAAATCAGAGA 1184

Db 245 TCCTGGAGACCAATGAGCTCTGATGAACGTCAGAAATTTATCAACAGAAATCAGAGA 304

QY 1185 AAGAGAGAGAGAAATCA-GAGAAAGCAGAGCCCATGAGCAGCGTGTAGAAAGC 1243

Db 305 AAGAGAGAGAGAAATCA-GAGAAAGCAGAGCCCATGAGCAGCGTGTAGAAAGC 364

QY 1244 CCCCACAGGCGCACACCAAGCTTACAGCCCAACCAACCACTTGTGTAGAG 1303

Db 365 CCCCACAGGCGCACACCAAGCTTACAGCCCAACCAACCACTTGTGTAGAG 424

QY 1304 AAGAGCTCTGTTCCAGATGAGCGCTGAGAAATGAGATTTTAAACAAAAATTCACAGC 1363

Db 425 AAGAGCTCTGTTCCAGATGAGCGCTGAGAAATGAGATTTTAAACAAAAATTCACAGC 484

QY 1364 TGCTACTTCTACAAAGACCTCACCCCAAGAGATCCCTGTGCAAGAGAGCTGCG 1423

Db 485 TGCTACTTCTACAAAGATCTCACCCCAAGAGATCCCTGTGCAAGAGAGCTGCG 544

QY 1424 CTGA-TTATGTGA-GAGAAATGAGACATTTATTTAGATCAAGACTTATTTGGCCC 1481

Db 545 CTGACTTATGTGAGCAGAGATTCAGACATTTATTTAGATCAAGACTTATTTGG-CC 603

QY 1482 AAGCGCTGGATACCAAGACCTGCAATGATCATCTGCAAGGAGGCTGTGAAATGGGA 1541

Db 604 AAGCGCTGGATACCAAGACCTGCAATGATCATCTGCAAGGAGGCTGTGAAATGGGA 663

QY 1542 GTCAACACTGCGCAGG-CAGGTGAAGAGCCTTGAGGAGAGGCGCAGCTGTATGAGGAC 1600

Db 664 GTCAACACTGCGCAGGCGCAGGTGAAGAGCCTTGAGGAGAGG-CAGCTGTATGAGGACA 722

QY 1601 GCTTC 1605

Db 723 GTTC 727

RESULT 14
B0003173/c 655 bp mRNA linear EST 26-MAR-2002
LOCUS B0003173
DEFINITION UI-H-E11-ayx-n-01-0-UI.s1 NCI CGAP_E11 Homo sapiens cDNA clone
IMAGE:5845056 3', mRNA sequence.
ACCESSION B0003173
VERSION B0003173.1 GI:19728073
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 655)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgabs-remail.nih.gov
COMMENT Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 32-117, >MER41BH1R/MER4-group
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers

FEATURES

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:02:31 ; Search time 291.391 Seconds
(without alignments)
8246.264 Million cell updates/sec

Title: US-09-996-617-1

Perfect score: 5444
Sequence: 1 gcccacgggcccgcgagagc.....aggataagaattaccac 5444

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165.8	3.0	1371	2 US-08-910-731-1	Sequence 1, Appl
2	165.8	3.0	1371	2 US-08-795-395-1	Sequence 1, Appl
3	163.6	3.0	1371	2 US-08-910-731-7	Sequence 7, Appl
4	163.6	3.0	1386	2 US-08-910-731-5	Sequence 5, Appl
5	156.2	2.9	1374	2 US-08-910-731-3	Sequence 3, Appl
6	156.2	2.9	1374	2 US-08-795-395-3	Sequence 3, Appl
7	115	2.1	585	4 US-09-340-620A-50	Sequence 50, Appl
8	115	2.1	740	4 US-09-340-620A-48	Sequence 48, Appl
9	106.2	2.0	579	4 US-09-340-620A-62	Sequence 62, Appl
10	106.2	2.0	777	4 US-09-340-620A-60	Sequence 60, Appl
11	62	1.1	7218	1 US-08-232-463-14	Sequence 14, Appl
12	59.4	1.1	257	4 US-09-016-434-208	Sequence 208, Appl
13	57.2	1.1	3116	4 US-09-904-615-43	Sequence 43, Appl
14	49	0.9	337	4 US-08-594-031-79	Sequence 79, Appl
15	48.6	0.9	1878	4 US-09-465-558-39	Sequence 39, Appl
16	48	0.9	1582	3 US-08-545-196B-10	Sequence 10, Appl
17	48	0.9	1582	3 US-08-545-196B-12	Sequence 12, Appl
18	47.2	0.9	17949	3 US-09-087-465-3	Sequence 3, Appl
19	46.2	0.8	599	3 US-09-328-111-147	Sequence 147, Appl
20	46.2	0.8	1441	3 US-08-821-994-63	Sequence 63, Appl
21	46.2	0.8	4141	4 US-09-245-281-42	Sequence 42, Appl
22	46.2	0.8	4141	4 US-08-207-359B-42	Sequence 42, Appl
23	46.2	0.8	4141	4 US-09-340-620A-42	Sequence 42, Appl
24	45.8	0.8	924	1 US-08-468-709B-1	Sequence 1, Appl
25	45.8	0.8	924	2 US-08-241-664B-1	Sequence 1, Appl
26	45.8	0.8	924	5 PCT-US93-03936-1	Sequence 1, Appl
27	45.8	0.8	1717	1 US-08-468-709B-6	Sequence 6, Appl

ALIGNMENTS

C 28	45.8	0.8	1717	2	US-08-241-664B-6	Sequence 6, Appl
C 29	45.8	0.8	1717	5	PCT-US93-03936-6	Sequence 6, Appl
C 30	45.6	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
C 31	45.2	0.8	289	3	US-09-007-005-17	Sequence 17, Appl
C 32	45.2	0.8	289	3	US-09-244-796-17	Sequence 17, Appl
C 33	45	0.8	1508	4	US-09-039-046-1	Sequence 1, Appl
C 34	45	0.8	11282	4	US-09-754-250-3	Sequence 3, Appl
C 35	44.4	0.8	1190	4	US-09-390-207-1	Sequence 1, Appl
C 36	44.4	0.8	3715	4	US-09-234-245-1	Sequence 1, Appl
C 37	43.8	0.8	467	2	US-08-841-349-18	Sequence 18, Appl
C 38	43.8	0.8	1868	1	US-08-658-883B-1	Sequence 1, Appl
C 39	43.8	0.8	1868	4	US-09-676-610B-26	Sequence 26, Appl
C 40	43.8	0.8	3437	3	US-08-860-339-17	Sequence 17, Appl
C 41	43.2	0.8	1342	4	US-09-489-847-89	Sequence 89, Appl
C 42	43.2	0.8	1534	1	US-08-300-903A-6	Sequence 6, Appl
C 43	43.2	0.8	1534	4	US-08-988-197-6	Sequence 6, Appl
C 44	42.8	0.8	1143	4	US-09-220-132-166	Sequence 166, Appl
C 45	42.8	0.8	1700	2	US-08-897-340-4	Sequence 4, Appl

RESULT 1
US-08-910-731-1
Sequence 1, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
ATTORNEY: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS

LOCATION: 1..1368
US-08-910-731-1

Query Match 3.0%; Score 165.8; DB 2; Length 1371;
Best Local Similarity 56.0%; Pred. No. 8.7e-34;
Matches 314; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

```
QY 2920 CTCTTCCTGCTGCTCAAGGTCACCAAGAACTGAAAGAGCTGACCTTAAGTGAACCTG 2979
DB 811 CTCTGCGTGTCTCCAGGCGCAAGAGACCTTAAGAGCTCACTGCGGGCAACAG 870
QY 2980 CTGAGCCACTCTGCTGAGTGAAGTCTTTGTAAGACCTTGAGAGCCCTGCTGCTCTG 3039
DB 871 CTGGGGGAGAGGGGCGCCCGGCTGCTGTGAGAGAGCTGTGTGAGCCCGGCTGCCAGCTG 930
QY 3040 GAGACCTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3099
DB 931 GAGTCCCTGTGGGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
QY 3100 GGGCTGAGAGCCAGCAAGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3159
DB 991 ATGCTGACCCAGAACAGCATCTCTGGAATTTGAGTGAAGCAAGAGCTGGGTGAC 1050
QY 3160 GCTGAGCCAAACACCTTTGCGAGAGCTGAGACAGCCGAGCTGAGCTGAGCTGAGCTG 3219
DB 1051 TCTGGCATCCAGAGAGCTGTGCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1110
QY 3220 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3279
DB 1111 TGTCTTGGGAGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1170
QY 3280 GCGAGCCCGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3339
DB 1171 GCGAACCGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1230
QY 3340 CGACTGCTGTGAGAGGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3399
DB 1231 CTGCACTGCTGAGAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1290
QY 3400 CAGCAACTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3459
DB 1291 GACACTTACTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1350
QY 3460 CAGCTGCTCATCTTTCAGCAGA 3480
DB 1351 GGCCTGAGGCTCATCTCTGA 1371
```

RESULT 2

US-08-795-395-1
Sequence 1, Application US/08795395

Patent No. 5965399

GENERAL INFORMATION:

APPLICANT: CHARTERED, DEB K.

APPLICANT: SHANDILYA, HARINI

TITLE OF INVENTION: Cloning and Expression of Rat Liver and
TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/795,395

FILING DATE: 04-FEB-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/024,057

FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: ESMOND, ROBERT W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.3440002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1371 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1368

US-08-795-395-1

Query Match 3.0%; Score 165.8; DB 2; Length 1371;
Best Local Similarity 56.0%; Pred. No. 8.7e-34;
Matches 314; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

```
QY 2920 CTCTTCCTGCTGCTCAAGGTCACCAAGAACTGAAAGAGCTGAGCTTAAGTGAACCTG 2979
DB 811 CTCTGCGTGTCTCCAGGCGCAAGAGACCTTAAGAGCTCACTGCGGGCAACAG 870
QY 2980 CTGAGCCACTCTGCTGAGTGAAGTCTTTGTAAGACCTTGAGAGCCCTGCTGCTCTG 3039
DB 871 CTGGGGGAGAGGGGCGCCCGGCTGCTGTGAGAGAGCTGTGTGAGCCCGGCTGCCAGCTG 930
QY 3040 GAGACCTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3099
DB 931 GAGTCCCTGTGGGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
QY 3100 GGGCTGAGAGCCAGCAAGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3159
DB 991 ATGCTGACCCAGAACAGCATCTCTGGAATTTGAGTGAAGCAAGAGCTGGGTGAC 1050
QY 3160 GCTGAGCCAAACACCTTTGCGAGAGCTGAGACAGCCGAGCTGAGCTGAGCTGAGCTG 3219
DB 1051 TCTGGCATCCAGAGAGCTGTGCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1110
QY 3220 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3279
DB 1111 TGTCTTGGGAGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1170
QY 3280 GCGAGCCCGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 3339
DB 1171 GCGAACCGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1230
QY 3340 CGACTGCTGTGAGAGGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3399
DB 1231 CTGCACTGCTGAGAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1290
QY 3400 CAGCAACTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3459
DB 1291 GACACTTACTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1350
QY 3460 CAGCTGCTCATCTTTCAGCAGA 3480
DB 1351 GGCCTGAGGCTCATCTCTGA 1371
```

RESULT 3

US-08-910-731-7

Sequence 7, Application US/08910731

Patent No. 5932440

GENERAL INFORMATION:

PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 740
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (54)...(638)
US-09-340-620A-48

Query Match 2.1%; Score 115; DB 4; Length 740;
Best Local Similarity 60.9%; Pred. No. 1.8e-20;
Matches 187; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 4476 CCCTCCAGCCGCGATAGCCCTTACCTCTGATGCCCCGAGTTGCTGACCTTGT 4535
DB 338 CCGCCAGCTGGATCCAGGCCCTCTCTAGTGGGAGCCAGCCAGGCTGACCTTAT 397
QY 4536 GAGCCGATTCGAGAGCAGCTGATAGCCGAGTGACATCGGTGAGTTGCTTGGACAA 4595
DB 398 AGACCGACACCGGCTGCGCTTATCGGAGGTCACAAAGTTAGTGGCTGCTGATGC 457
QY 4596 ACAGCATGACAGATGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4655
DB 458 TCTGTACGGAGAGTCTTACGAGTGGAGTACAGGAGGAGGAGGAGGAGGAGGAG 517
QY 4656 GCCAGCAGATGCGAAGCTGTTCAGCTTGAAGCAGTCTCTGAGCCGAGGAGGAG 4715
DB 518 CCCAGCAAGATGCGAAGCTTTCAGTTTCAACCCAGCCTGGAAGTGGAGGAGGAG 577
QY 4716 TGACCTTACCAAGCCTTGAAGAGAGCCATCTCTACCTCATTTATGAACTTGGAGAA 4775
DB 578 CTGTCTCTCCAGGCTTAAGAGAGTCCAGTCTCTACTGTGAGAGCCTGGAGCGAG 637
QY 4776 GGGCAGC 4782
DB 638 CTGAGGC 644

RESULT 9

US-09-340-620A-62
Sequence 62, Application US/09340620A
Patent No. 6482933
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/340,620A
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 62
LENGTH: 579
TYPE: DNA
ORGANISM: Mus musculus
US-09-340-620A-62

Query Match 2.0%; Score 106.2; DB 4; Length 579;
Best Local Similarity 64.4%; Pred. No. 3.3e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 4528 CACTTGTGACAGATGAGAGCAGTGTAGAGCCGAGTGAATCGGTGAGGTTGTC 4587
DB 331 CACTTGTGACAGATGAGAGCAGTGTAGAGCCGAGTGAATCGGTGAGGTTGTC 390
QY 4588 TTGACAACTGATGAGCAGTGTGAGCCAGAGCAGTGAAGAGGTTGCTGAG 4647
DB 391 CTGATGCTTTCATGAGTGTGCTGACTGAAGACAGTACAGGAGTTCGTGAGAG 450
QY 4648 AACAGAGGCCAGATGAGAGCTGTTGAGCTTGAAGCAGTTCGGAGCGGAAG 4707
DB 451 ACCAGCAGCAAGATGAGAGGCTTTCAGCTTTCATCTGGAACCTGACC 510
QY 4708 TGCAAGATGACTTACCAAGCCTTGAAGAGAGCCATCTCACTCATTTATGAATC 4767
DB 511 TGCAAGATGACTTCCCTCCAGGCTTGAAGAAATCATCCCTACTTGGTGAATGACCTG 570
QY 4768 TGGGAGA 4774
DB 571 GAGCAGA 577

RESULT 10

US-09-340-620A-60
Sequence 60, Application US/09340620A
Patent No. 6482933
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/340,620A
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 777
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (89)...(667)
US-09-340-620A-60

Query Match 2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 4e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 4528 CACTTGTGACAGATGAGAGCAGTGTAGAGCCGAGTGAATCGGTGAGGTTGTC 4587
DB 419 CACTTGTGACAGATGAGAGCAGTGTAGAGCCGAGTGAATCGGTGAGGTTGTC 478
QY 4588 TTGACAACTGATGAGCAGTGTGAGCCAGAGCAGTGAAGAGGTTGCTGAG 4647
DB 479 CTGATGCTTTCATGAGTGTGCTGACTGAAGACAGTACAGGAGTTCGTGAGAG 538
QY 4648 AACAGAGGCCAGATGAGAGCTGTTGAGCTTGAAGCAGTTCGGAGCGGAAG 4707
DB 539 ACCAGCAGCAAGATGAGAGGCTTTCAGCTTTCATCTGGAACCTGACC 598
QY 4708 TGCAAGATGACTTACCAAGCCTTGAAGAGAGCCATCTCACTCATTTATGAATC 4767
DB 599 TGCAAGATGACTTCCCTCCAGGCTTGAAGAAATCATCCCTACTTGGTGAATGACCTG 658

QY 4768 TGGAGA 4774
Db 659 GAGCAGA 665

RESULT 11

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pl-F1s
US-08-232-463-14

Query Match 1.1%; Score 62; DB 1; Length 7218;
Best Local Similarity 2.8%; Pred. No. 8.5e-06;

Matches 11; Conservative 230; Mismatches 145; Indels 0; Gaps 0;

QY 23 GAAGAACTCGGAGCAGCCCGGCTCTGAGCTTGAGAGCCATTCCT 82
Db 1052 GAGGAGCTTCGCAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1111
QY 83 GCTCTGGGCTCTCCACCCACCTCTTCAAGCTTGACGTCAGATCTCAG 142
Db 1112 YY 1171
QY 143 GAGTCGAGACCCGAGAGGAGAGATCTGAGAAACAGACAGTGGCTGCCAC 202
Db 1172 YY 1231
QY 203 ACCCATCTCCGCTACCAATCTCCCTCACCTCACCTCCTGCTGAGCCTGAGAC 262
Matches 11; Conservative 230; Mismatches 145; Indels 0; Gaps 0;

Db 1232 YY 1291
QY 263 CCATCCAGACCTCCCTATCAGCTGACTTTCAGATCTTGAGGCCCTTGAGCT 322
Db 1292 YY 1351
QY 323 CTTCCCTCCCTGCTTCTTCACTCCCTCCCTATGAGGCTGATGAGTCC 382
Db 1352 YY 1411
QY 383 TGGATTTATMAACTGGGTTCCGAA 408
Db 1412 YYYYYYYYYYYYYYYYYYYYYYYYGTA 1437

RESULT 12

US-09-016-434-208
Sequence 208, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SINTBST01
CLONE: 1499408
US-09-016-434-208

Query Match 1.1%; Score 59.4; DB 4; Length 257;
Best Local Similarity 57.5%; Pred. No. 5.2e-06;

Matches 122; Conservative 0; Mismatches 89; Indels 1; Gaps 1;

QY 1987 TATTTCAAGATTTTCAAGATGAAGCAAGCAATTGAGCTTTAGCTGCTGAA 2046
Db 2 TATTTCAAGATTTTCCGAGATGAGAGGAGCGAGCGCCCTTACCTTCGTAG 61
QY 2047 TCAACAAAGAGCTGGGCGCTGTGTCTGTGCTGCTGCTGCTGCTGCTGCT 2106
Matches 122; Conservative 0; Mismatches 89; Indels 1; Gaps 1;

TYPE: DNA
ORGANISM: Trilicium aestivum
US-09-465-558-39

Query Match 0.9%; Score 48.6; DB 4; Length 1878;
Best Local Similarity 52.8%; Pred. No. 0.013;
Matches 105; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY	5237	AGATCTACAGAGCAAGCCACAGAGTAAAGTGAAGAACTTATTCAGAAATAAAG	5296
Db	1672	AGATCCATGATGAGCCGCCACATATGTAAGATCCAGCAATTTACGAATATGGG	1731
QY	5297	AGTATCAGAGCTCTTTAGAAATTTGTAGCAGGCTTCCAGTTTACCAGAAACCCC	5356
Db	1732	AACTTGTCAATAGTTTCTTAATTCAGGCACATCTGTAAGATCTTCGCTGATATGCTA	1791
QY	5357	TATTAATTAATAATTTTACTTAATTTAAGAAATTAACAAAAAGAAAAAT	5416
Db	1792	TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1851
QY	5417	GAAAAATAAGAAATTAAGAA	5435
Db	1852	AAAAAAAAAAAAAAAAAAAAA	1870

Search completed: January 29, 2004, 13:11:27
Job time : 293.391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 35.197 Seconds

(Without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 1 MAGAMGRLACYLEFLKKE.....HLIMELMKSGKGLPLSS 1429

Sequence: 1 MAGAMGRLACYLEFLKKE.....HLIMELMKSGKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5969	79.2	1192	T17255	hypothetical prote
2	882	11.7	1111	A59000	matric protein [imp
3	465.5	6.2	483	S27880	Nasopressin recept
4	361.5	4.8	456	S20597	ribonuclease inhib
5	358	4.8	461	A31856	ribonuclease-angio
6	354.5	4.7	456	A31857	ribonuclease inhib
7	345	4.6	1130	A48843	MHC class II trans
8	206	2.7	1075	T31668	hypothetical prote
9	202	2.7	1004	T31665	hypothetical prote
10	198	2.6	1232	A55478	neuronal apoptosis
11	196.5	2.6	1447	T42628	neuronal apoptosis
12	180	2.4	533	T52063	ran GTPase-activat
13	166.5	2.2	1120	UC7765	mitotic spindle as
14	159.5	2.1	506	A45841	T-complex-associat
15	155.5	2.1	1199	T47442	disease resistance
16	153.5	2.0	1131	T96662	hypothetical prote
17	152	2.0	1027	T46296	hypothetical prote
18	152	2.0	3511	A59295	unconventional myo
19	151	2.0	1221	T52347	disease resistance
20	150.5	2.0	1220	T48928	disease resistance
21	149	2.0	631	C89243	protein F28C1.3 [i
22	149	2.0	631	T21471	hypothetical prote
23	148	2.0	1217	T52348	disease resistance
24	147.5	2.0	793	S54772	mammary gland fact
25	147.5	2.0	1214	T47438	disease resistance
26	147	2.0	1017	H96663	hypothetical prote
27	147	2.0	1784	T10532	gag-pol polyprotei
28	143.5	1.9	4302	A38971	polycystic kidney
29	142	1.9	1215	T00364	hypothetical prote

30	142	1.9	1839	1	OYBYK	adenylate cyclase
31	139.5	1.9	545	2	T52068	ran GTPase-activat
32	139.5	1.9	1189	2	T52346	disease resistance
33	138.5	1.8	1031	2	G96663	hypothetical prote
34	138	1.8	1010	2	T36383	probable large ATP
35	137	1.8	1226	2	T45788	transcription disea
36	136.5	1.8	794	2	G02317	adherence factor T
37	136.5	1.8	3335	2	H81702	probable resistanc
38	135.5	1.8	1422	2	B71437	smoothened - human
39	135	1.8	915	2	T09575	P-glycoprotein - S
40	135	1.8	1724	2	T18343	hypothetical prote
41	134	1.8	889	2	H96606	hypothetical prote
42	133	1.8	1164	2	H71436	disease resistance
43	133	1.8	1301	2	D85188	hypothetical prote
44	133	1.8	1453	2	G96613	ATP-dependent prot
45	132.5	1.8	805	2	AH2731	

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFZ58601822.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17255

R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18722

A:Accession: T17255

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1192 <KO>

A:Cross-references: EMBL:AL117470

A:Experimental source: adult uterus; clone DKFZ58601822

C:Genetics:

A>Note: DKFZ58601822.1

Query Match	79.2%	Score 5969	DB 2	Length 1192
Best Local Similarity	96.2%	Pred. No. 0		
Matches 1147	Conservative 1	Mismatches 0	Indels 44	Gaps 1
QY	282	LLLRPHRPSQDPLVKSWSVDYVENRGLIEIRDLFGPGIDTQEPRIIVILQAGAGIGKS	341	
DB	1	LLLRPHRPSQDPLVKSWSVDYVENRGLIEIRDLFGPGIDTQEPRIIVILQAGAGIGKS	60	
QY	342	TLARQKEAMRGQLYGDREFOHVFYFSCRELAOSKVSLAELIGKDGTAAPAPRIOLISR	401	
DB	61	TLARQKEAMRGQLYGDREFOHVFYFSCRELAOSKVSLAELIGKDGTAAPAPRIOLISR	120	
QY	402	PERLLFTLDGVDSEGWLYQSPSSSELCLHWSQPOPADALQSLQKTLIPENAFILITARTT	461	
DB	121	PERLLFTLDGVDSEGWLYQSPSSSELCLHWSQPOPADALQSLQKTLIPENAFILITARTT	180	
QY	462	ALQNLIPSLQARVETLGSSESRKRYFRFYFDEQAIRAPLVSNKELMLCLVPW	521	
DB	181	ALQNLIPSLQARVETLGSSESRKRYFRFYFDEQAIRAPLVSNKELMLCLVPW	240	
QY	522	VSWLACTCLMQMRKRLTITSTKTTTLCILHYAOLAQLOAPLQPORLDCSLAEGIMQ	581	
DB	241	VSWLACTCLMQMRKRLTITSTKTTTLCILHYAOLAQLOAPLQPORLDCSLAEGIMQ	300	
QY	582	KKTLPSPDRLRKGLDGAISTFLKMGILQEHPIPLSYFTHLCQSFPAAMSVLEDEK	641	
DB	301	KKTLPSPDRLRKGLDGAISTFLKMGILQEHPIPLSYFTHLCQSFPAAMSVLEDEK	360	
QY	642	GRGHSNCIIDLKTLTAAGHGLFGASTTRFLGLSLSDGGERMENIFHRLSOGNNIM	701	
DB	361	GRGHSNCIIDLKTLTAAGHGLFGASTTRFLGLSLSDGGERMENIFHRLSOGNNIM	420	
QY	702	QWVSLQLLQPHLSLHCLCYETRNKTFLOVAAHFEENGMCEVETMELLVCTFCIKFS	761	

Db 421 QWPSLQLLQPHSLBSLHCLYETRNKTFLTQVMAHEEMGCVETDMELLVCTFCIKFS 480
 QY 762 RHVKUQLIEGRORHSPTMVLFRWVPTDVAWQILFSLKVTEMLKELDLSGNSLS 821
 Db 481 RHVKUQLIEGRORHSPTMVLFRWVPTDVAWQILFSLKVTEMLKELDLSGNSLS 540
 QY 822 HSAVKSICKLRPRCLLETTLRLAGCGLTAEDCKDLAFGRANQTLTELDLSRNVLTDA 881
 Db 541 HSAVKSICKLRPRCLLETTLRLAGCGLTAEDCKDLAFGRANQTLTELDLSRNVLTDA 600
 QY 882 AKHLCORLQPSCKLQRLQVLSGCLTSDCCODLASVLSASPSLKELDLQNNLDVGVRL 941
 Db 601 AKHLCORLQPSCKLQRLQVLSGCLTSDCCODLASVLSASPSLKELDLQNNLDVGVRL 660
 QY 942 LCEGLRPACKLRLGLDQTLTSDENKQELRALQEKRPOLLIFSRKPSVMTFTEGLDTG 1001
 Db 661 LCEGLRPACKLRLGLDQTLTSDENKQELRALQEKRPOLLIFSRKPSVMTFTEGLDTG 720
 QY 1002 EMENSTSLRQRLGSRASRASHVAQANLKLDVSKTFPIAEIABESSPEVVPVELLCVPS 1061
 Db 721 EMENSTSLRQRLGSRASRASHVAQANLKLDVSKTFPIAEIABESSPEVVPVELLCVPS 780
 QY 1062 PASQGLHTKPLGTDDPFMGPTGVATEVVDKEKILYRVAFPVAGSYRMPNTGLCFVYRE 1121
 Db 781 PASQGLHTKPLGTDDPFMGPTGVATEVVDKEKILYRVAFPVAGSYRMPNTGLCFVYRE 840
 QY 1122 AVTVEIEFCVWDQFLGEINQHSMMVAVGPLLDIKABEVAEVAHLPHFVALQSGHVDTS 1181
 Db 841 AVTVEIEFCVWDQFLGEINQHSMMVAVGPLLDIKABEVAEVAHLPHFVALQSGHVDTS 900
 QY 1182 FQVAHFEBEGBMLLEKPARVLEHITVLENBSPFLGVLKMTIHNAIRPTSTVVLVYHRV 1241
 Db 901 FQVAHFEBEGBMLLEKPARVLEHITVLENBSPFLGVLKMTIHNAIRPTSTVVLVYHRV 960
 QY 1242 HPEBEVFLHLLIPSDCSIR----- 1260
 Db 961 HPEBEVFLHLLIPSDCSIRKAIDLEMKPQVRIHKRPPLTPLYMGCRVTVSSGSGMKLE 1020
 QY 1261 ---KELELCYRSPGEQDLFSEFVYGHLSGIRLQVQDKQDETLVWEALVYKPDLMPTTL 1317
 Db 1021 ILPKLELELCYRSPGEQDLFSEFVYGHLSGIRLQVQDKQDETLVWEALVYKPDLMPTTL 1080
 QY 1318 IIPARIAVPSPLAPOLHFDVQYREQLIARVTSVEVVLDKHGOVLSQCYRVALENT 1377
 Db 1081 IIPARIAVPSPLAPOLHFDVQYREQLIARVTSVEVVLDKHGOVLSQCYRVALENT 1140
 QY 1378 RPSQMRKLFSLQSWDRCKDGLYOALKEITHPHLIMELMEKSGKGLPLSS 1429
 Db 1141 RPSQMRKLFSLQSWDRCKDGLYOALKEITHPHLIMELMEKSGKGLPLSS 1192

RESULT 2

A59000

mater protein (imported) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C/Accession: A59000

R/Tong, Z.B.; Nelson, L.M.

Endocrinology 140, 3720-3726, 1999

A>Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ov

A/Reference number: A59000; MUID:99360614; PMID:10433232

A/Accession: A59000

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1111 <KUR>

A/Cross-references: GB:AF074018; NID:G5802697; PIDR:AAD51762.1; PID:G5802698

C/Genetics:

A/Genes: Mater

Query Match 11.7%; Score 882; DB 2; Length 1111;
 Best Local Similarity 26.1%; Pred. No. 2,4e-49;
 Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

QY 162 SSPDHSPEQESBNAP---TSTAVLGSWSPPOPS---LAPR---EQEAPGTQWPLD 209
 Db 28 TSPENDSKSIQKQGEQEOQTESGTWG---PPEKSKAIIKARGLSEOKSSTSPS 83
 QY 210 EMSGIYITEIRERERKSEKGRPMAAVVGTTPQANTSTQPHHHPMEPVRSLSGTWVW 269
 Db 84 NYSRAILKQSGSEBEVQASERK---WTSPEKSKSIQKQGEQEOQTESGTWG--- 133
 QY 270 KNEDFNQKFTOLLRLRPHRSGDPLVYKRWDPDYENRGHLE----- 313
 Db 134 KEED---EYTE-----ADKONGD-LQDYKAHYIAKDTSDVLDHYSPENK 175
 QY 314 -IRDLFGPELDQEPPIVYLQGAAGIKSTLARQVKAENGQGLYDRQHYFYSCRL 372
 Db 176 LLSDAKFPYQKTFQPTIIILHGRPGVKSALRSIYLAGAOGKL-F-QKMSFYIFFSVRBI 234
 QY 373 AGSKVYSLAEILGKQATAPAPRIOTLSPERLLFLDGDVBGWTLOBPSSBLCLHWGQ 432
 Db 235 KTEKSSLLQILAKCEPDSDWLVTKIWSQPERLLFYIDGLDNDVSLQHDMLTSRDKD 294
 QY 433 POPADALGSLGKTLPEASPLITARTALQNLIPSLBOARWEVLAGSESSRKEYFYR 492
 Db 295 EQPIYILAMWSILRKALLPQSFILITTRNTGLEKSMVVSPLYLVEGLSASRRQGLVIE 354
 QY 493 YPTDERQAIRATRLVYKSKELMALCLVPVWSLACTCLMQKREKTLTSTKTT--- 548
 Db 355 NTSNESDRIQVFRHSLIENHQLPDQCAPSVCSLVCEALOLQKLGKRCCTLPCQTLTGLYA 414
 QY 549 TICIHLV-----AQLQAQPLRQRLDLSLAEGIMQKTLFSPRDLRKHGDAIIST 603
 Db 415 TLVFHQLTLKRSQSLSEBOQTLVLGCMMAABGVWTKRSFYDDDLKQYSKESEILA 474
 QY 604 PLKMGILQIE--HPIPLSYFHLCPQEFPAWSYVLEDEKGRKH-----SNCIIDLEK 655
 Db 475 LFMNITLLQVGNHSEGCYFHSLSLQDFPALUYVLEGLBEMNQHCPIENGOSIMEVR 534
 QY 656 TLEAVIGHLFGASTTRPLGLISDGEREMENIFRCRL--SQGRYLMQWVBSQLLLP 713
 Db 535 TDDT---RLDMK--RFLGLMKNQDILKTLEVEYPIFVBOQLQHWV---SLIAQ 585
 QY 714 -----HSLSLHCLYETRNKTFLTQVMAHEEMGCVETDMELLVCTFCIKFSRHVK 766
 Db 586 VNGTSMQTLDAFYCLFESQDEFPVQALKRQGVWLLINQMDLVSSYCLKHQCNLKA 645
 QY 767 LQI---IEGRORHSPTMVL- FRWVPTDVAWQILFSLKVTRNLKELDLSGNSLS 821
 Db 646 IRVDIRDLSTVNTLELCVYTVVQETQCKPLIMEMWGNFCVLSIRNLKELDLSGNSLS 705
 QY 822 HSAVKS----- 828
 Db 706 QRAMKILCLELRNOSGRLOKTLFKSAENVVSGLKHLMKLLFSNQNLKYLNIGNTPMKDDM 765
 QY 829 ---CKTLRRPRCLLETLR----- 843
 Db 766 KLACEALHKPKSVETLRDSCETITIGEMISTILLTRKLCLSLAKRNVGVSMISL 825
 QY 844 -----LAGCGLTAEDCKDLAFGRANQTLTELDLSFNVLTDAGAKHLCOR 888
 Db 826 GNALSSMCLQKLLIDNGCLTPASCHLLVSAIFSQNLTHLCLSNNSIGTEGVQQLCOF 885
 QY 889 LQPSCKLOR----- 898
 Db 886 LRNPBCALQRLILNMCNIVDAVYGLAMRLANNYKLTHTLSLTMPNPGDAMKLLCEALKE 945
 QY 899 -----LQVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLCEGLRHAP 951
 Db 946 PCYVIGLELTVQQLQNCEDLACMITTKLHKSLDNNMLGDKGVITLCEGLKQSSS 1005
 QY 952 KLRIGLDQTLTSDENKQELRALQEKRPOL 981
 Db 1006 SLRRIGLAGCKLTSNCCALSLAISCNPHL 1035

RESULT 3

S27880

Naopressin receptor - rat
C/Species: Rattus norvegicus (Norway rat)C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C/Accession: S27880

R/Herrera, V.L.; Ruiz-Opazo, N.

submitted to the EMBL Data Library, February 1992

A/Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.

A/Reference number: S27880

A/Accession: S27880

A/Molecule type: mRNA

A/Residues: 1-483 <HER>

A/Cross-references: EMBL:M65183; NID:9202805; PIDD:AAA03623.1; PID:9202806

C/Genetics:

A/Gene: AVP

Query Match 6.2%; Score 465.5; DB 2; Length 483;

Best Local Similarity 30.6%; Pred. No. 1.3e-22;

Matches 161; Conservative 74; Mismatches 186; Indels 105; Gaps 18;

QY 534 MKRKKLTLSKTTTLCVLAQALQ-AQPLGP---QLRLCSLAEGIMQKTLFSP 588

Db 1 MEICRDLRSRTSKTTSTVYLLFITSMLKSAGTNGPRVQGLRMLCRLAREGILKIQAPSE 60

QY 589 DDLR--KHGDGAIIST-FLKM---GIDRHPILPISFTHLCFQEPFAAMSVLEDEK 641

Db 61 KDRLRLKTLKQSGVQTMFLSKKELPGVLE---TVVTYQFIDQSFQEFLLASTYLDAB- 116

QY 642 GRGKSNICIIDLEKTLFAYGHLGPG--ASTTRFLGLLSDGEGREMENTPHC---RLS 695

Db 117 --GAPGASAGSVQMLMSDA--GLRGHLALTTRFLGLLSTERRIDIGNHFGCVPRVK 172

QY 696 QGRNLQWVPSLQLLLP-----HSLSLHCLY 723

Db 173 Q--DTRWV--QQSQPKVATGAKKDEKDEAESEEESEELNPGELVCLY 227

QY 724 ETNKKTLTVQMAHFEEMGCV---TDNELVCTGICIKSRVKKQLQI-----E 771

Db 228 ETQEDDVRAALSLPE--WYLERVRLTRMDLEVLSYVCCPDGQALRLVSCGLVAKE 285

QY 772 GRRHRSWSPTMVYLFPMVPTDAVQWQILFSLVLTNKLKELDLSGSLSHSAVSKLCKT 831

Db 286 KKKKKKSF-----NRLKSGSGTGQPPASLIRPLCEA 318

QY 832 LRPRRCLETLRLAGCGLTAECDKDLAFGLRANQTLTELDLSFNVLTDAAGKLCORLRQ 891

Db 319 MITQCCGLSITLISHCKLPAYVCRDLSEALKVAPSLARELGLQRLTEAGIRLLSQGLAW 378

QY 892 PSCRLQRLQVSCGLTSDCCODLASVTSAPSLKELDLQONNLDVGVRLICGLRHPAC 951

Db 379 PKCVQQLRIQMPGL-OBVHYLYIVAQOSEVLTTLTDLSCGQLPGTVVEPLCSALKHPKC 437

QY 952 KTLRLGIDQTLTSDENKQELRALQEKPOLIFSRKRPVWTPNG 997

Db 438 GLKTLSTSVLELTENPLRELDQAVTKLPDLAIHSLKLTGHPQLKG 483

RESULT 4

S20597

ribonuclease inhibitor - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000

C/Accession: S20597

R/Kawamoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.

Biochim. Biophys. Acta 1129, 335-338, 1992

A/Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution

A/Reference number: S20597; MUID:92162755; PMID:1556887

A/Accession: S20597

A/Molecule type: mRNA

A/Residues: 1-456 <KMW>

A/Cross-references: EMBL:X64528; NID:957670; PIDD:CAA44386.1; PID:957671

C/Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 4.8%; Score 361.5; DB 2; Length 456;

Best Local Similarity 35.5%; Pred. No. 7.9e-16;

Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;

QY 792 VTDAYWQILFSLVLRNKLKELDLSGSLSHSAVSKLCKTLRRRCLETLRLAGCGLTA 851

Db 206 ITSANCKDLCDVVASKASLDELDSGSLKGLNTGTAALCSGLLPSCRLTLMWDCVTA 265

QY 852 EDCKDLAFGRANQTLTELDLSFNVLTDAAGKHLCORLRQPSCKLQRLQVSCGLT--- 908

Db 266 EGCKDLQVLRANQSLKELSLAGNELKDBAQQLCSLELPQCQLSELMWTKSLTAAAC 325

QY 909 -----DC-----CQDL 914

Db 326 PHFCSVLTKNSSLPFLQMSNPLGDSGVBLCKALGYPTVLRVLMGDCVDTSGCSL 385

QY 915 ASVLSAPSLKELDLQONNLDVGVRLICGLRHPACKLRLGIDQTLTSDENKQELRAL 974

Db 386 ATVLNRSRLRELDLSNCKMGDNGVLIQLLESLKQPSCLIQQLVLYDIYWTDEVDQRL 445

QY 975 EOEKPOLIFS 985

Db 446 EERRPLRITS 456

RESULT 5

A31858

ribonuclease-angiotensin inhibitor - human

N/Alternate names: ribonuclease inhibitor, placental

C/Species: Homo sapiens (man)

C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000

C/Accession: A31858; S02012; S23933; S48636; T47188

R/lee, P.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.

Biochemistry 27, 8545-8553, 1988

A/Title: Primary structure of human placental ribonuclease inhibitor.

A/Reference number: A31858; MUID:89118269; PMID:3219362

A/Accession: A31858

A/Molecule type: mRNA

A/Residues: 1-461 <LEB>

A/Cross-references: GB:M22414; NID:9186260; PIDD:AAA59130.1; PID:9307040

R/Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.

EMBO J. 7, 4151-4156, 1988

A/Title: The primary structure of human ribonuclease/angiotensin inhibitor (RAI) disclosed

A/Reference number: S02012; MUID:89210799; PMID:3243277

A/Accession: S02012

A/Molecule type: mRNA

A/Residues: 1-422, SE', 425-461 <SCH>

A/Cross-references: EMBL:X13973; NID:935843; PIDD:CAA32151.1; PID:935844

A/Note: part of this sequence, including the carboxyl end of the mature protein, was cont

R/Czeval-Thieffry, I.; Cotterill, S.; Schuller, E.

Biochim. Biophys. Acta 1122, 107-112, 1992

A/Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibiti

A/Reference number: S23933; MUID:92358217; PMID:1633192

A/Accession: S23933

A/Molecule type: protein

A/Residues: 174-195 <CRB>

R/Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.

Arch. Biochem. Biophys. 312, 421-428, 1994

A/Title: Purification and characterization of human brain ribonuclease inhibitor.

A/Reference number: S48636; MUID:94311593; PMID:8037455

A/Accession: S48636

A/Status: preliminary

A/Molecule type: protein

A/Residues: 2-14 <NAD>

R/Pouska, A.; Melnreuther, R.; Mewes, H.W.; Well, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A/Accession: T47188

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-461 <AAA>

A; Cross-references: EMBL:AL161967
 A; Experimental source: adult testis; clone DKFZp434K249
 C; Genes: GDB:RNH
 A; Cross-references: GDB:125274; OMIM:173320
 A; Map position: 11p15.5-11p15.5
 A; Note: DKFZp434K249.1
 C; Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 F; 34-43/Region: leucine-rich 57-residue repeats
 C; Keywords: blocked amino end; duplication
 F; 57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F; 114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F; 174-195/Region: inhibitory
 F; 228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F; 285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F; 399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

Query Match 4.8%; Score 358; DB 2; Length 461;
 Best Local Similarity 37.9%; Pred. No. 1,4e-15;
 Matches 86; Conservative 30; Mismatches 91; Indels 20; Gaps 3;

QY 768 QLEIGHQRHSTWSPVTWVLFVWVPTDAVWQI-----LPSVLKVTNLEKELDLSGN 818
 DB 246 ELCPGLIHPSRRLT-----WI-----WECITAKGCGDLCRVLRAKESLKEISLAGN 294
 QY 819 SLSSHAVKSLCKTLRRPCLLETLRLAGCGLTAECDKDLAFGLRANQTLTELDLSPNVLT 878
 DB 295 ELDEGARRLCELTLEBGCQLESIMWVSCFTAACSHFSVLAQNRFLELDISNNRLE 354
 QY 879 DAGAKHLQRLRPSCKLQRLQVSCGLTSDCCODLASVLSAPSLKEDLQNMIDVVG 938
 DB 355 DAVVRELCOGLQPGSVLRVLMADCDVSSCSLSATLNLNHSRELDLSNNCLGDAG 414
 QY 939 VRLICEGLRHAPACKLIRLGLDQTLTSDENKQELRALEQKRPQLIFPS 985
 DB 415 ILQLVSVNQPGCLLEGLVLYDYTWSEWEDRLQALEKDPRLRIVS 461

RESULT 6
 A31857
 A; ribonuclease inhibitor, hepatic - pig
 C; Species: Sus scrofa domestica (domestic pig)
 C; Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
 C; Accession: A31857; A35830
 R; Hofsteenge, J.; Kieffer, B.; Mathies, R.; Hemmings, B.A.; Stone, S.R.
 Biochemistry 27, 8537-8544, 1988
 A; Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals th
 A; Reference number: A31857; MUID:89118268; PMID:3219361
 A; Accession: A31857
 A; Molecule type: Protein
 A; Residues: 1-456 <HOF>
 R; Vincentini, A.M.; Kieffer, B.; Mathies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
 Biochemistry 29, 8827-8834, 1990
 A; Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea
 A; Reference number: A35830; MUID:91104783; PMID:2271559
 A; Accession: A35830
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 82-456 <VIC>
 A; Cross-references: GB:M58700; NID:g164638; PIDN:AA63448.1; PID:g164639; GB:J02925
 C; Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 C; Keywords: liver
 F; 280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR>

Query Match 4.7%; Score 354.5; DB 2; Length 456;
 Best Local Similarity 34.3%; Pred. No. 2.3e-15;
 Matches 109; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

QY 752 LVCTFCIKFSRHVVKQLIEGRHSTWSPVTWVLFVWVPTDAVWQI-----LPSVLKVTNLEK 811
 DB 127 LICEGLIDPQCHLEKQL-----EYCRITAAACBPASVLRATRAK 168
 QY 812 ELDSGNSLSHAVKSLCKTLRRPCLLETLRLAGCGLTAECDKDLAFGLRANQ-TLEL 870

DB 169 ELTVSNNDIGEAGVILQGLADLSACQLETLNLENCGLTPACKQLC-GIVASQASLREL 227
 QY 871 DLSFNVLTDAAGAKHLQRLRPSCKLQRLQVSCGLTSDCCODLASVLSAPSLKEDLDQ 930
 DB 228 DLGSGNLGDAGIAELCPGLISPARSLKTLTLMWECITAGSGCDLRCVLAOKETLKEISLA 287
 QY 931 QNNLDVGVRLICEGLRHAPACK-----LIRLGLDQTT 962
 DB 288 GKRKGEGARRLICESLQPGCQLESIMWVSCSLTAACCHVSLMTQNGHLELDLSNKR 347
 QY 963 LDEKRGEL-RALDEKQQLIFSRKRSVMTPTGDLTGMSNS-TSGIKRQRLGSEPA 1020
 DB 348 LQDSGIQELCOALSQPGTLRLVLC-----LQDCEVTNSGCSL-----A 386
 QY 1021 ASHVAQAMKILDVVS 1035
 DB 387 SLLNLNRSIRELDLS 401

RESULT 7
 A48843
 MHC class II transactivator - human
 N/Alternate names: CIITA
 C; Species: Homo sapiens (man)
 C; Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
 C; Accession: A48843
 R; Steimle, V.; Otten, L.A.; Zufferey, M.; Mach, B.
 Cell 75, 135-146, 1993
 A; Title: Complementation cloning of an MHC class II transactivator mutated in hereditary
 A; Reference number: A48843; MUID:94006336; PMID:8402893
 A; Accession: A48843
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-1130 <STE>
 A; Cross-references: GB:X74301; NID:g414112; PIDN:CA52354.1; PID:g414113

Query Match 4.6%; Score 345; DB 2; Length 1130;
 Best Local Similarity 23.2%; Pred. No. 4e-14;
 Matches 256; Conservative 134; Mismatches 430; Indels 282; Gaps 49;

QY 13 LEFLKKE-----ELKEFOLLANKHSSSSSETPA-----QPEKTSQMEVAS 55
 DB 111 LELSKDITKRIHGPDEVIGESWEMPAEVQSKQKPPFELPADLKNKPAEPVYVTS 170
 QY 56 YLVAQYGEORAWDLALHTWQWGLRSLC-----AQAGEAGHSPPSPSEPHLGSPSP 111
 DB 171 LTVGPVSD-----CSTLPCLPALFWQEPASGMRLKXTDQIPMPSSSSLSGLNLP 223
 QY 112 TSTVAMPMIHELPACTQGSERRV-----LRQLPTSGRRMBISASLTY 157
 DB 224 EGIQCFVPTISTLPHGLWQISEAGTGVSSIFVHGSEVPOASQVPPPSG-----FTV 274
 QY 158 QALPSPDHESPPQSPNAPTAVLWGSQSPPOSLAR-----BOEAPGTOMPLDETSG 213
 DB 275 HGLPFSBDR--PGSTSPFAPSATDL-----PSMBEPALSRAMTTEKTPPTQCP--AAG 325
 QY 214 IYVTEIRERERKSEKGRPPMAAVVGTPTQATSLQPHHPWEPVSRESLCTSWPKND 273
 DB 326 -----EVSNK-LPKMPEPV--EQFYSIQDTGGA-EPAPDGI----- 359
 QY 274 FNQKFTQLLLQRPHPRSQDPLVKR--SWPDYVEN--RGHLIEIRDLEFGGLDIOEPRI 329
 DB 360 ----LVEVDLVQARLERSSSKSLERELAPDWAERLQAGGLAEVLLAAKEHRRPRETV 415
 QY 330 VILQGAAGIGKSTLARQVKEAMGRGQLYDROHVFYFSCRELAQ-SKVYSIAELIGKDG 388
 DB 416 IAVLGKAGGKSYWAGAVSRANACGR--PYDFFPVSVCCLNRRGDAYGLDILLFSIG 473
 QY 389 TATPAPI-----RQLSRPERLLFTLDGVDEPGVNLQEPSSSLCLHWSQOPAD---- 437
 DB 474 ----PQPLVADAVFVSHILKRPVILLIDAPFE-----LEADGFLHSTCGAPAPAPCGL 525

QY 438 -ALGSLIGLTILPEASFLITARTTALQNLIPSELA-RWVEVLGFSRSSRKEVEYRYET 495
 DB 526 RGLAAGLFQKILKCECTILLTRPRG--RLVQSLKALALFELSGFSQOAYMYRFE 583
 QY 496 D-----ERQAIRAFRLVKSNEKLMALCLVPVSWLACTCMOQMKREKLTITSTKTTTLC 551
 DB 584 SSGMTEHOD-RALTILRRPRLILSHSHSPILCRANVCQSLBALLEGEDAKLPS-TLTLGLY 641
 QY 552 LHYLAQALQAPLPQRLDLCSLAAE-GIMOKTTL---FSPBDLRKGLDGAIIISTLK 606
 DB 642 VGLGRALDPPG-ALAEELAKLAMELGRHOSITLQEDQFSPADVR-----TWAMA 691
 QY 607 MGLIOEHP-----IPLSY-SFTHICF-QEFAAMSIVLEDE-----KGRGHSICI 650
 DB 692 KGLVQHPRAASELAPSPFLQCFGLALMTALSGEIKDKELPOLYALTLPKKRPYDQWL 751
 QY 651 IDLEKTEAVYGHGLFGASTRFLGLLSDG-----EREM 686
 DB 752 EGVRPFLA-----GLIPPARCLGALLGPSAASVDKQVLAHYLKRLOPGLIRARQL 806
 QY 687 ENTHFC-----RLSGRNLQWVP-----SLQILLQPHSIE--- 717
 DB 807 LEHLHCAHEAEAGIWOHVQELPGRSLPLGTRLTPDAHYLGKALEAGQDFSIDLRST 866
 QY 718 -----SLHCLVETRKKTFLTYMAHFEEMQCVETDM-----ELLVCTFCIKF 760
 DB 867 GICPSGLSGLVGLSCV--TRPRAALSDVALWESLRQGETKLQAAEKKTTIEPFKXS 924
 QY 761 SRHYKUL-QLIEGRHSTWSPPTWVL-----FRWVPVTDAY-WQILFSLVKYTR 808
 DB 925 LKQVEDLQKLVQTRSSSEDTAGELPAVADKLKEFALGPVSGPQAPFLVAILTAFS 984
 QY 809 NLKELD--LSGNSLSHSAVSKLCTLRPRCLLETTLAAGCGLTAEDCKOLAFGLRANQ 865
 DB 985 SLQHLDLDAISENKIGEGVQSLATF--POL-----K 1015
 QY 866 TLTELDLSPNLTAGAKHLGRLRQPSCKLQRLQVLSGCLTSDCCODLASVLSASPSLK 925
 DB 1016 SLETLNLSQNNITDLAGAKLALPSLALSLRSLYNNCICDVGASLAVLDPWVSLR 1075
 QY 926 ELDLQONLDDVGVRLCEGLR 947
 DB 1076 VMDVQYNKFTAAQAQQLAASLR 1097

 RESULT 8
 T31668
 hypothetical protein Cosl.5 - sea equit (Ciona intestinalis)
 C:Species: Ciona intestinalis
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T31668
 R:Bird, A.P.; Clark, V.; Jones, S.J.; Lettgeb, S.; Lennard, N.; Tweedie, S.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z21050
 A:Accession: T31668
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1075 <BIR>
 A:Cross-references: EMBL:Z80904; PIDN:CA802589.1
 C:Genetics:
 A:introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1
 Query Match 2.7%; Score 206; DB 2; Length 1075;
 Best Local Similarity 18.5%; Pred. No. 4.8e-05;
 Matches 209; Conservative 152; Mismatches 347; Indels 422; Gaps 49;

QY 395 IROILSRPERLLFLIDVDPERGWTLOEBSSELCLHWSOPADALGSLGKTLIPENSF 454
 DB 156 ITPLHNNPNAIF-FDGLDEASTNEFARIPRICLIDGSKSEVD-IMKNLFWLTLPPKAKI 213
 QY 455 LITARTTALQNLIPSELAQWVEVLGFSRSSRKEVEYRYFDEQAIRAPFLVKSNEKLM 514
 DB 214 VVITSLHQMVKLHDYDPTSLFEVLGLLEAKNNLGLQLCSEKTPALK--KILDOQPLA 271
 QY 515 ALCLVPVSWLACTCMOQMKREKLTITS---KITTTTCLHYLAQALQAPLPQRLD 570
 DB 272 HLCTPLINFLIVFCILSN--EGSDIKMTQVLIISWTRFVLSHLKGRVPLDKYGAENVX 330
 QY 571 LCSIAAGIWKQKTLFSPBDLRKGLDGAISTF-----LKNGILOEHPILSYS 620
 DB 331 LARLAAYKGLQQRKLVFEKTDIFDVKLADENVTFPHYVDKSGSIRMKILGN--KRSY- 387
 QY 621 FTHCFQSPFPAAMSIVL---FSTYREFQKPIFKDAQMKRVGFMFGICNPAYKOLKLVF 637
 DB 388 FTHLWQSFYAVYLMFLVSTYREFQKPIFKDAQMKRVGFMFGICNPAYKOLKLVF 447
 QY 638 -----EDEK-----GRG----- 644
 DB 448 ATMIKDYBEKKELAVPMMSLSMARGBDLIRRGMLHEVNDSSKKFEDYLPVGLMD 507
 QY 645 --KSNCTIIDLEKLEAY-----GHLGFGASTRPLG-- 676
 DB 508 APKILSEVKDLYVALKFTPKRLDLSYETTTTEVLETLRGVHGT-TTITTFEVINNI 566
 QY 677 -----LISGEEREMNI-----FHGCLSGRNLQWVPSPQLLOLPHSLSL 719
 DB 567 EMKDSLELLHLHDAMEELFIRVTNLSPYMERLSNAIN--QNSNKIQVLVLRH--KL 621
 QY 720 HCLVETRKKTFLTYMAHFEEM--GMCVETDMELVCTFCIKSRHYKULQIEGRQR 776
 DB 622 HDYDVK--YIACGLNLSILYMWGTDISSDQ-----CSVLKQAIQPLPSIQ-- 666
 QY 777 STWSPVWVLFRWVPVTDAYWQILFSLVKYTRN-----LKELDLSGNSLSHSAVSKLCK 830
 DB 667 -----VHQLYPDILSTYLVANRNISWPFSEVLDVPEALN----- 702
 QY 831 TLRRPCLLETTLAAGCGLTAEDCKDLAFGLRANQTLTELDLSPNLTAGAKHLGRLR 890
 DB 703 -----LNGUL-----KDSKQPRRNEELCSQDVSVAPLTVQVYNYHYC--N 740
 QY 891 QPSCKLQRLQVLSGCLTSDCCODLASVLSASPSLKELDLQONLDDV----- 937
 DB 741 APYHQVRF-----TQAFILF--MINSLNHNGLDQYEKIMRLDSEYLE 784
 QY 938 GVRLLCEGLRHPA-----CKLI-----RLGLD 959
 DB 785 TPKLQCE---PASLTPKEQVITIKLQTHVVLDEKETIRCKLHSDGITWETRYTKLEFS 840
 QY 960 QTTLS-----DEMROELRALDEQEKPOLLFSSRKPSVMTPTSGLD--TGEMS 1004
 DB 841 DNFTSPQNTFSSWKVIFEFLEKTLAPLYKVKLSLYLQANAGVWYKCMNLDVATGE-- 898
 QY 1005 NSTSLKROLGSERAASHVAQANTKLIDVSKTPIPIAIEBSSPEVPVELCVPS-- 1061
 DB 899 ---NELKENHTQGTW--IPLSNDLILCEK-----HENAEVYINI--IPSGKI 941
 QY 1062 -PASQ-----GDLHTKPLGTDDDFQWFGTGVATEVVDKEXKLYRVHFPVAGSY 1108
 DB 942 IPANQUNNSYCCNKKFVVDKIS--TNE-----VHLIKAK-----CGSF 978
 QY 1109 RWPNTGLCFVWBEAVTVIERCVWDQ--FLGEINPQHSMMVAVAGPLLDIK 1155
 DB 979 RW-DDDFCF-----PLPITHAVSDQOQOPTOPSIN--TYITGPTTTIR 1019

 RESULT 9
 T31665
 hypothetical protein Cosl.4 - sea equit (Ciona intestinalis)

```

C:\Species: Ciona intestinalis
C:\Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:\Accession: T31665
R:\Bird, A.P., Clark, V., Jones, S.J., Leilegeb, S., Lennard, N., Tweedie, S.
Submitted to the EMBL Data Library, October 1996
A:\Accession number: 221050
A:\Accession: T31665
A:\Status: preliminary; translated from GB/EMBL/DBJ
A:\Molecule type: DNA
A:\Residues: 1-1004 >A1R>
A:\Cross-references: EMBL:280904; PIDN:CAB02586.1
C:\Genetics:
A:\Introns: 30/1; 539/3; 593/3; 662/3; 728/3; 769/3; 817/2; 881/2; 960/1

Query Match          2.7%; Score 202; DB 2; Length 1004;
Best Local Similarity 17.7%; Pred. No. 7.9e-05;
Matches 191; Conservative 162; Mismatches 373; Indels 352; Gaps 40;

244 AHTSLQPHHHHPREVSRESLSCTSWKXKDEPFQKFTQLLQRPHPRODPLVYKSWEDY 303
   |||:::|||||
   9 AHAAKKHQHP-----EPAVNVQPLV-----30
Db

304 VEENRGHLIEIRDLFGPGIDT-----QEPRIYILQAGAGIKSTAR 345
31 -----TIDQLFDKALENAOCTTEOQSAEYAKYIERHANTYVWVGPHGGKTIILK 81
Db

346 QVKEAMRGQDYGRFOHVFYFSCRELAQSKVSLAELIGKDTA-----TPAPIRQ 397
82 MNVNOQLKHELPRDT-EYIFPIYADIDPNKEMTLLEFLTNSRKVNVYTESKALTTF 140
Db

398 ILSRPKRLFLIDGYDEFGWVLQDESSSELCHMSOPAPADLLSGLLKTIIPKASFLIT 457
141 LHNPNVVAIF-FDGIQDEASLKELVGYSQICXLDKSKRPVD-IMKNLFNALLPKAKIVVT 198
Db

458 ARTALQNLPLSLBQARWVYVLSGFSSSRKKEYFYVYFDERQATRAPFLYSNKEIMLALC 517
199 STPDEMFLDQCYRPTSLFEVLGFLBKKNNLIGTOLQSEKIPALF-KILDQPNYLAHC 256
Db

518 LVPWVSWLACTCLM-----QOMKREKTLTSTKTTTLTCLAHYLAQALQAPLGPOLDLCS 573
257 YLPINFLIIVPCLLSNEGSDIKTMQVLAIFMSTRVESH-LKGVLPIDKYGAEWVTLAC 315
Db

574 LAASGIWQKTLFSPDDLRKGLDGAIISTF-----LKNGILOENHPILYSSTIYH 623
316 LAYKGLQQRKLVFEKTDIPDDVLADEWNTNFEHTYVDISSGIRIKILBGN--KASY-PTH 372
Db

624 LCFQEPFAFNVULDEDEKGRGHSNCIIDLEKTLAEGHIGFGASTRPFLLGLSDSGE 683
373 HIMQEFYAAV-----YLMLEFYSYREF 393
Db

684 REMERIFACRLSQGNLMOVPSQLTL---QPHSLSLAHCL---YETRAKTELT 732
394 EQLKTIFFD-----TQMSVVYCFMGICNPFAHYKDLKIIIPATMIKYDEKKKFLFS 445
Db

733 QVNAHFEMGACVETDMELLVCTFCIKFSRHVYKQLQILEGROHSTWSPTWVVLFRWVPV 792
446 VVM---BSLSAKSEHDL-----IRRF---GWLHEYNDETCKKFCQCLPVL 484
Db

793 TDAWQQLIFSVLYKTVTRNLKEILDSNSLSHSAVNSLCTIARPROLLETLLAGGGLTAE 852
485 -----GLMVGVPKHLPEYKDLVYAKSTYTKPKHKLRSNWTTTTEVLTLLRGHIGTTTY 539
Db

853 DCKDLAFGLRANQUTLELDSFNVLTLDGAKHLCQRLRQPSCKLQRLQVLVSCGLTSDCCQ 912
540 ITRFVANNIEKKDSIMELL---LHLD-----MEKLAFDVNTLMSIME 581
Db

913 DLASVLSASPLKEILD---QQNNLDVGVRLLCGLRHPAKCLIRLGIDOTTYSDEKRO 969
582 SLNSAINORSKTIQDILWHQILNDDDVYLAGLIG-----NISRLNMSHYIISD--- 632
Db

970 ELRLLEQKPKQLLFSRRKPSVMTTEGLDTPQEMNSNTSSLKQGLGSEBPAASHVAQANL 1022
633 QCRVYLKQALBQD-----PST-----647

```

Oy	1030	KLLDVSKIPIPIAEIAESSEGVVVELLCVPSPASQDHTKPLGTDDPMGPTGPVATE	1089
Dd	648	--OVHQY-----DLIST-YLVNWRPILRDEMT-----	674
Oy	1090	VVDENKULYRVHFPVAGSYEMPTGLCFVRREAVTVEIEECWDOFLGEINPOHSNWAG	1149
Dd	675	-----SVYFHH-----DQQFESSSKCMIGSRGGKLEVGCC-----	704
Oy	1150	PULDITKAEPQAVE---AVHLPHFVALQGCHVDTSLPQMAFKERGMILEPARVELH-HI	1205
Dd	705	---ELVYPFGALBKDVVEIKLTASLSIESFELETPTLOC---ELASLTLKQVTYIKLQTHV	758
Oy	1206	VLENPSYSPICVLLKMIMNALRFIPVTSVVLLVHRHYRPEEVTFLH---	YLISP-CSI 1259
Dd	759	VLDKET-----IRCKYTLVTRIV---TTTVHMKGKMLNHIDICSI	796
 RESULT 10 A55478 neuronal apoptosis inhibitory protein - human			
N	Alternates names:	NAIP	
C	Species:	Homo sapiens (man)	
C	Date:	05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001	
C	Accession:	A55478	
R	ROY, N. de Jongdevan, M.S.; McLean, M.; Shuler, G.; Yaraghi, Z.; Farahani, R.; Baird, S.		
d	T.O., J. Mahadevan, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; Mackenzie, A.		
A	Title:	The gene for neuronal apoptosis inhibitory protein is partially deleted in indi	
A	Reference number:	A55478; PMID:95112344; PMID:7813013	
A	Accession:	A55478	
A	Molecule type:	mRNA	
A	Residues:	1-1232 <ROY>	
A	Cross-references:	GB:U19251	
C	Genetics:		
A	Gene:	GDB:SMA@; SMA	
A	Cross-references:	GDB:120378; OMIM:600354; OMIM:253300	
A	Map position:	Sg12.2-Sg13	
A	Keywords:	apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prot	
F	9/94-110/Domains:	transmembrane #status predicted <TM1>	
F	4/79-477/Region:	nucleotide-binding motif A (#-loop)	
F	4/79-456/Domains:	transmembrane #status predicted <TM2>	
F	4/76/Binding site:	ATP (lys) #status predicted	
F	6/18,632,823,923,1035/Binding site:	carbohydrate (Asn) (covalent) #status predicted	
 Query Match 2.6% Score 198; DB 2; Length 1232; Beet Local Similarity 20.8%; Pred. No. 0.0002; Matches 158; Conservative 108; Mismatches 282; Indels 212; Gaps 34;			
Oy	292	QDPVKSMPDYVENRGHILIRDLFGPGDLQOPRIVLQGAQIGKSTLARQYEAW	351
Dd	450	QEPLV-----LPEVYG-NLNS---VMCEBASAQKVLAKTAFLW	487
Oy	352	GRQL-YGDRFOHFVFSCRELAAQS--VSIAELIGDKGTAPAPIROISR-BERLI	406
Dd	488	ASGCCPLNFQVLYFYLSSTRPDDEIASIIDQLLEKGSYTEMCNRNIIOQLNQVL	547
Oy	407	FILDGVDEPWVQEBSESLCHMSQPQPADALLGSLIGKITLIPSAFLTARTTALQNL	466
Dd	548	FLADY-----KEIC---SIPQ---YIGKIQNHLISRCCLLAVINRADRI	589
Oy	467	IPSLQARWEVLGFSSSRKEYRYFTDERQAIRFL-VSNKELMALCIVPVWSWL	525
Dd	590	RRVLETI--LEIQAFPPYNVCILRKLFISHNMRLKRMFWPEKQSGLOIGTPFLVA	647
Oy	526	ACTLMOQ-----MKREKULTSTKTITTLCHYLAOLAOAPLGQLRDLS	573
Dd	648	ICAHMFPYPPDSPBDVAVAFKSYMERLSTNKAT-----AEILKATVSSCGE	694
Oy	574	LAAEGVIQKTKLSPDVLRRGHGDGAIISFLMGAILQHBPILASTYSFILCOEPPAAM	633
Dd	695	LAKGFSCCFERNDDDLAEAGVDEEDLTMCMSKFTAQRKLPPFRFLSPAFOEFLAGH	754

QY 634 SYVLEDEKRGKHSNC-IIDLEK-----TLEAVG-----IHGL-----FGASTTRFLIGL 677
 DB 755 RLIELBDBDROEDDGLYLHLKQINS PMMTVASINPNFNVSLSTAGPKYSHLHL 814
 QY 678 LSDGEREMENIFHCRLSQGRNLQWVPSLQLLQPHSLSLHCLYETRNKTFITQVMAH 737
 DB 815 V--DNKESLENI-----SENDYLNKQBEISLQWQ-----LLRGLMQCPQAYFSMSEH 862
 QY 738 FEEMGMCVETMELLVCTFCIKFSAHVAKKLQILBR-----QHRSTWSPTWV 784
 DB 863 L--LVLAKTAYQSNVTAAQSPF-----VLQFLQGRITLTLGALNLQYFEDHPEBSILRS 915
 QY 785 VLF-----RWYVPTDAYQVILFVVL--KYTRRLKELDLSGNLS 821
 DB 916 IHFSIRGKTSPPRAHFSVLETCFDSQVPTIDQYASAFEPMMNERNLAE-----K 967
 QY 822 HSAYSLCKTLRRPRCLLETIRLACGGLTAEDCKDLAFG--LRANQ--TLTELDSLFWVL 877
 DB 968 EDNVKSYMDQORRA-----SPDLSTGYWKLSPKQYKIPCLQLEVDVNDI 1009
 QY 878 TDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQONNLDV 937
 DB 1010 DVVG-----QPMLEILMTVFSASQRI-ELHILNHS----- 1037
 QY 938 GVRLLCBGLRHPACKLIRLGLDQTLTSDMRQELRALQOE 977
 DB 1038 --RGFIESIR-PALELSKASVTKCSIS--KLELSAAQOE 1071

RESULT 11

T42628
 neuronal apoptosis inhibitory protein 2 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
 C:Accession: T42628
 R:Yaraghi, Z.; Diez, E.; Gros, P.; Mackenzie, A.
 Mamm. Genome 10, 761-763, 1999
 A:Title: cDNA cloning and the 5' genomic organization of Naip2, a candidate gene for murine
 A:Reference number: Z22179; MIMD:99315342; PMID:10384056
 A:Accession: T42628
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1447 <VAR>
 A:Cross-references: EMBL:AF102871; NID:G3860228; PID:G3860229; PIDN:AACT3002.1
 C:Genetics: Naip2

Query Match 2.6%; Score 196.5; DB 2; Length 1447;
 Best Local Similarity 19.7%; Pred. No. 0.00032;
 Matches 172; Conservative 138; Mismatches 303; Indels 261; Gaps 43;

QY 329 IVILQGAAGIKSTLQGVKAWGRGOL-YGDRFGHVFYFSC-----RELAQSKVSLA 381
 DB 509 VMCEVGEAGSGKTYPLKRIAFILMAGCCPLNRFQLVYFLSLSTPQGLAK--IIICA 565
 QY 382 ELIGKDGATAPARQILSR-PERLLFTLDGVDEGWLQEPSSBLCLHWGQOPADALL 440
 DB 566 QLTGGGCGISSEVCSSIIQLOLQHVLPILD-DYSGLA-----SLPQ--AL 607
 QY 441 GSLGKTLPEASFLITARTALQVLSLBOAWEVLGSESSRKEYFYRYFDERQA 500
 DB 608 HTLITKYLSTCLLIIVHTNKVRGIRPYLDTS--LEIKFPFYTVSVLRLGTFSHDMR 665
 QY 501 IRAF-RLVSNKELMALCLVPSVSWLACT-----CLMQQMKKEXL 541
 DB 666 VRKFINYGFHBLGHIKTPLEVAACVCTDWPKNPSDQPFQDVALFKAYMQLSLKHGA 725
 QY 542 LTKSTTTTLCIHYLAQALQAPLQPLDLSLAEGIWOKKTLFSPDLKRGKID-GAI 600
 DB 726 -----AKPLQATVSSCGQLALTGLFSSCFEPNSONLBAVGEDEBE 766
 QY 601 ISTPLKMGILQEHPIPLSYSPFHLGFOEFPA-----MSYVLEDEKRGKHSNCIIDLE 654

DB 767 LTTCLMSKFTQGRARPV-YRIGPLFOEFLAAVRLTELLSSDRQEDDGLIX--YLROIN 823
 QY 655 KTLBAYGIGHLF-----GASTTRFLIGLSDGEREMENIFHCRLSQGRNLQW 703
 DB 824 SPLKMSIYHFLKLVYSSHPSSKAAFTVSHLLQVLEKE-SLEN-----MSENDYDKL 877
 QY 704 VPSQLQLQPHSLSLHCLYETRNKTFITQVMAHFEEMGMCVETMELLVCTFCIKFSRH 763
 DB 878 HPEALLW-----ISCLNGLMQSPSSFSFLTSEN-----LRLCLNFAHE 917
 QY 764 VKK-----LQIEGR-----QHRSTWS-PTWVLFK-----WV----- 790
 DB 918 SNTVAACSPVLQRLRGRITDLKVLSLQYFMDHETILLKSLKSLNGNMVQRIIDSL 977
 QY 791 -----PVDAWQILFS-VLKYTRNLKEL-----DISGN--SLSHSAVSL 828
 DB 978 IEKSEKQVQRPPTIDQYALAPQINQVQKNSKKHIIKYEYDMKHQIPLNISTGYKLS 1037
 QY 829 CKTLRRPRCLLETIRLACGGLTAEDCKDLAFGLRANQTLTEL-----DISFNVLTD 879
 DB 1038 PKPYKIRK-DE-VQVNTVG-----PADQALLQVLMVEVFSASQSTIEFRILDS 1081
 QY 880 AGAKHLQRLRQPSCKLQRLQVSCGLTS-DCCODLASVLSASPSLKELDLQ--NNLDP- 936
 DB 1082 SG--FLESIR-PALELSKASVTKCSMSRLESLREDQKLLTLPTLQLEVSETNQLPDQ 1137
 QY 937 -----VGTRLLC-----EGIRHPACKLIRLGLDQTLTSDMRQEL 971
 DB 1138 LFNHLHKFLGKELCVRLDSKPDVLSVPGEPNLIHH--MEKLSIRSTESD--LSKL 1191
 QY 972 RALFOEKQQLILF-----SRKRP--SVMTPTGDLGEMSN----- 1005
 DB 1192 VKLIQNSBNLHVFLKCNFLSNCEPLMTVLAASCKKLEIEFSRGCFEAMTFVNIIPNVF 1251
 QY 1006 -STSLKQRQLGSERAASHVAQANIKLIDVSKIF 1038
 DB 1252 LKIIMLRQCPDDETSKFRQALGSLRNLEKLF 1285

RESULT 12

T52063
 ran Gase-activating protein (imported) - alfalfa
 C:Species: Medicago sativa (alfalfa)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52063
 R:Pay, A.; Nick, P.; Nagy, F.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25929
 A:Accession: T52063
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-533 <PAY>
 A:Cross-references: EMBL:AF19528.1
 A:Experimental source: subspecies x variat, cultivar A2

Query Match 2.4%; Score 180; DB 2; Length 533;
 Best Local Similarity 21.1%; Pred. No. 0.0008;
 Matches 110; Conservative 76; Mismatches 188; Indels 148; Gaps 16;

QY 544 SKTTTLCIHYLAQALQAPLQPLDLSLAEGIWOKKTLFSPDDLKRGKIGLCAIST 603
 DB 89 AKSSSKMLLEVLPKQPRKEENGEL-----ISRGDAVAVTVFDISGRRAPIDQEASE 143
 QY 604 FLKMGILQEHPIPLSYSPFHLGFOEFPA-----GASTTRFLIGLSDGEREMENIFHCRLSQ 657
 DB 144 LKK-----PLMGPNSTFKICFENRSGDADAIVYEMWLSTIDQLKEVDLSPIAGR 196
 QY 658 --EAYGHGLFGASTTFL-----GLSDGEREMENIFHCRLSQ 696
 DB 197 EAEIEVWNIFSSALBFAVIRYLNLSNNAMGEKGVARFALLKQONDLIEELVYLNWDGISE 256
 QY 697 --GNLMQWVPSLQLLQPH-----SLESJHCLVETRNKTFILT 732

Db 257 EAKAVAEELPSTKLVLFPHNMVGTDEGAFAIAEVKSPALIEDPRC---SSTRVGE 313
 QY 733 QVNAHEEMGCVETDMEELLVCTFCIKFSRHVKKLOIEGRQHRSTWPTMVVLFRRVAV 792
 Db 314 GGVALAEALGACT-----HLKTLDT----- 333
 QY 793 TDVAWQI-----LSVLYKTRNLKEDLDSGNSLSHSAVKSLCTLRPRCLLETTLRLAGC 847
 Db 334 RDNMFGEAGVALSKVLPVPADLTETIYSTLNIEDDGAALANALAKESAPSLTLDMAGN 393
 QY 848 GLTAEDKDLAFGLRAQNTLTEDLSPFVLTDAAGAKHLCORLRQPCCKLRLQVSGCLT 907
 Db 394 DITKAVASVAECSSKQFLTKMLSENELKDSGA-----GL- 430
 QY 908 SDCCQDLASVLSASPSLKELDLQONNLDVGVRLLCBG-LRHAPCKLIRLGLDQTLTSD 966
 Db 431 -----ISKALBGRGQSEVDLSTNLITWSGAKLLAAVAVQKPGFKL--LNINANPISDE 482
 QY 967 MRQELRALBQEKPOLLFSSRKPSVMTPTGDLTGEMSNST 1008
 Db 483 GIDELKDIIFKNSPDMT-----GLDENDPGEEDVDEAEADDS 519

RESULT 13

JCT765
 mitotic spindle associated protein, MAP126 - human
 C/Species: Homo sapiens (man)
 C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C/Accession: JCT765
 R/Chang: M.S.; Huang, C.J.; Chen, M.L.; Chen, S.T.; Fan, C.C.; Chu, J.M.; Lin, W.C.; Yart
 Biochem. Biophys. Res. Commun. 287, 116-121, 2001
 A/Title: Cloning and characterization of hMAP126, a new member of mitotic spindle-associ-
 A/Reference number: JCT765; PMID:11549262
 A/Contents: Testis
 A/Accession: JCT765
 A/Molecule type: mRNA
 A/Residues: 1-1120 <CHA>
 A/Cross-references: GB:AF345347
 C/Comment: This protein, as a nuclear protein and a new member of kinesin family, which
 C/Genetics: and dynamic regulation of mitotic spindles and may provide an outward force to push the
 A/Map position: 17q11-12

Query Match 2.2%; Score 166.5; DB 2; Length 1120;
 Best Local Similarity 18.8%; Pred. No. 0.02;
 Matches 217; Conservative 136; Mismatches 391; Indels 411; Gaps 50;
 QY 148 WR--EIASLLYQALPSSPDHSPQSPPNAPTAVLGSWGPSPQSLAPR-----EQE 200
 Db 2 WRVKKLSLSPSPQTKPSMRPLRLILQPGALTTSGK-RSPACSLTPTSLCKGLQE 60
 QY 201 APGTQWPLD-----ETSGIYYTE-----IRERESEKGRPMAAVGTTPQARTSLQ 249
 Db 61 GSNNSSPVDVFNKRTLDLSSHFSSSKMLTQCHHEDEQDLDTIPQISSTRKTSSEAVID 120
 QY 250 PHHHPWEPVSRESLCTWPKNEDFNQKFTQLLLQRPHPSPQDPLVKRSPDYVEENRG 309
 Db 121 PLG-----NYMVKTIVLVPSPPLGQQQM----- 143
 QY 310 HLIETRLPFGRLDTQ-EPRVILQGAAGIKSTLAQVKEAMRGQLYGDPRQHVYPS 366
 Db 144 -----IFEARLDVTMTSISLNGP--LRTDDLVEEV-----APCMGDRF----- 182
 QY 369 CRBLAOSKVSLAELIKDGTATPAPRIQILSRPERLLFLIDGVDEPGWVLOEPSBLC- 427
 Db 183 -----SEVAVSE-----KPIFQ-----ESSHLLESPRPPCS 211
 QY 428 --LHWSQOPADALLGSLGKTLIPBASFLITATTL-ONLISLSLQA-----RVVEV 478
 Db 212 EQLHCSKES-----LSRTEAVREDLVPSBSNAFLPSSVLW--- 247
 QY 479 LGFSESERKEYFYRYFTDERQAI-----RAFR-----LVXSNKEMLALC 517

Db 248 LSPSTALADFRVNHVDPSEIEVHEGMEERERFPTHKESSETDQALVSSVEDILSTC 307
 QY 518 LVP-----WVSMIACGLAQMOQRKKEKTLTTSKTTT 549
 Db 308 LTPNLVMESEGAEPRAVEDVGRILISDTPESWMSPLA-----WLEKGNITS 353
 QY 550 LCHLYAQAQLQAQPLQRLDLC-----SLAEGIWQKTLFSPD-----DLRKHGL 596
 Db 354 VMLENLRQSL--LPSMLRDAIGTTPFTSCVGTW-----PTSPAPQEKSTNTSGTGL 405
 QY 597 DGAIIITFLKMGILQHPRIPLSYSPFH-----LCPOEFPAANSYVLEDEKGR- 643
 Db 406 VGTKSTSETQQLCGRPDPDTALSRHDEDLNLSLVIYEF--LSRQLRWKSQLAVP 462
 QY 644 -----GKHSNCIDLEKTELEAYGIGHLFGASTTRPFLGLISDEGEREMENIFH 691
 Db 463 HPTQDSSTQDTSHSGITNKLOHLESH-----EMQG-- 495
 QY 692 CRLSGRNIMQ-WPSELQILLQPHSLSLCHLYETRNKFTLTQ----- 733
 Db 496 -ALQQRNVWQSWVLISKELI--SLHLSLHLEEDKTTVNQESRRAEVLVCCFPDLK 551
 QY 734 -----VNAHEEMGCVETDMEELLVCTFCIKFSRHVKLO--LIEGRHRS 777
 Db 552 KIRAKLOSLKAEEREARHREEMALRGKDAEIVLEAFCAHASQRIQLEODLASMEFRG 611
 QY 778 TWSPTWVLFRRVPTDVAWQILFVLYKTRNLKEDLDSGNSLSHSAVKSLCTLRPRC 837
 Db 612 LKMDAQTL---VGLAKQELVQGVVSLTSLIQ--DMSMOJDTYTWTL--LSRSRQ 664
 QY 838 LLETTLRAGCGLTAECDKDLAFGLRANQITLEDLSPFVLTDAAG-AKHLQRLRQPSCK 895
 Db 665 LTKELT-----VKSQAQLQERDVAIEEKQVSRVLEQVSAOLEBECKGQ 707
 QY 896 LQRLQVSGSLTSDCCQDLASVLSASPSLKELDLQONN-LDDVGV-R-LICE----- 944
 Db 708 TEQLELENRLADLPRAQLQITLNMDSQLTELOSQTHCAQDLAMKDELQTLQSNREQ 767
 QY 945 -----GLRHACKLIRLGLDQTLTSDMRQELRALE-----QE 977
 Db 768 AAGCVBEMALKMQLQELQ---QAVLAKVAVDLKTELEFADQENQVHLIEGYECQL 824
 QY 978 KPOLLIFSRP-----KQSVMTPTGDLTGEMSNSTSLIKRQLSERRAASHVAQANLK 1030
 Db 825 KTLIEVLRESLQCNENIKQVENI/LAKIASTIADNOEDLEKTRQVSOXKGLTTEQ--- 880
 QY 1031 LLDVSKIFPIAELAESSEPRVAVPELLCVSPSPASQGDHLTKPLGTDDEFKGPVATEV 1090
 Db 881 -LQSLTFLQTKKKEKTEQETLTLSTACPT-----QEHPLPNDRTF--LGSILITAV 929
 QY 1091 VDKEKNLYRVHPVAGS 1107
 Db 930 ADEPESTPY--PLDGS 944

RESULT 14

A45841
 T-complex-associated-testes-expressed-1 protein - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
 C/Accession: A45841
 R/Sarvetnick, N.; Tsai, J.Y.; Fox, H.; Pilder, S.H.; Silver, L.
 Immunogenetics 31, 283-284, 1990
 A/Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
 A/Reference number: A45841
 A/Accession: A45841
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-506 <SAR>
 A/Cross-references: GB:M28821; NID:g201909; PIDN:AAA40406.1; PID:g201910

Query Match 2.1%; Score 159.5; DB 2; Length 506;

Best Local Similarity 20.5%; Pred. No. 0.016;
Matches 117; Conservative 77; Mismatches 194; Indels 183; Gaps 25;

QY 1 I R A F R L V Y S N S E L M A L C L P W V S W L A C T C L M Q M K R K E K L T L S K T T T L C L H Y L A Q L Q 560
Db 55 V R R R R I S E D A E N S I A V L P L L ----- T E L I Q I H V K N F Q 89
QY 561 A Q P L G P Q R L D C S L A E G I W Q K T L F S - P D L R K H G L D G A I I S T L K K G I I Q E H P I P L S Y 619
Db 90 N N P I - - - - - E K O P L E - - H Q R K V L S N L P P E L ----- P L T V 117
QY 620 S F I H L C F O E F P A A S V Y L E D E K G R G K S N C I I D E K T L E A V G I - - - - - H G L F G A S T T R E L 674
Db 118 T - - - - - A N L I D D E - - N Y W R K C C I - - - - - K R M S V C H V S R H G - - G S W K R M F F 153
QY 675 L G L L S D E G E R E M E N I F H - - - - - C R L S O G R - N I M Q W V P S I Q L L L O P H S 715
Db 154 - - - - - E R H L E N L K L F I P G T T D P V I L D L P L C R N V Y A R I H V D Q L P F V R M - - - P T P 202
QY 716 L - - - - - E S I H C L Y E T R N K T F L Q W A H F E M - - - - - G M C V E L D M 749
Db 203 L Q G E S D S G S E G S E B E K D H Y Q L O T - - - - - L V G K L H E E L D L V Y G V D Q G N N F E N L 257
QY 750 E L A V T C F I K S R S R V Y K K L Q L I E G R O H R S T W S P T W V L P R V Y - - P V T A Y W O I L P S V L K V T 807
Db 258 F L F Y R R C Y S I A A T I K A C H T L K - - - - - I F K L T R S K V D D K A I L R S L I D H 303
QY 808 R N I K E L D I S G N S L S H S A V K S L C K T L R R P C L L E T L R L A G C G L T R A E D C K D L A F G L R A N Q T L 867
Db 304 P A L E E L D I S H N L I G D R G A R A A K L L S H R - - L R V I L A N N Q L Q P G A Q S L A H L A H N T N L 361
QY 868 T E D L S E R V L T D A K A K H L C Q R L R O P S C T Q L Q L V S C G L T S D C C D D A S V I S A S P S L K E L 927
Db 362 V F L N L R N C I E D E G Q A I A H A L E T N K C - L S V L H L G G N K L S E P T A T L S O M L T V T T V L S L 420
QY 928 D L Q Q N N L D V G V R L L C E G L R P A C K L I R L G D Q T T L S D E M Q E R A L E O E K P O L L - - I F S 985
Db 421 N L S G N H I G O D G K Q L L E G I S D - - - - - N K I L E P D L R - L S D V S Q E S E Y I L G V L H 468
QY 986 R R K P S V M T P T E G L D T G E M S N S T S L K R O R L G 1016
Db 469 A N R E A R Q R T - - L N P G H P S S P T N C T E N S V G 497

RESULT 15
T47442
disease resistance protein homlog - Arabidopsis thaliana
N/Alternate names: protein T18B22.70
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T47442
R/Jordan, N.; Bangerter, S.; Wiedemann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke, K.
Submitted to the Protein Sequence Database, February 2000
A/Reference number: 224467
A/Accession: T47442
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1119 <DOR>
A/Cross-references: EMBL:AL138652
A/Experimental source: cultivar Columbia; BAC clone T18B22
C/Genetics:
A/Map position: 3
A/Intons: 57/1; 238/2; 613/3; 713/3
A/Note: T18B22.70

Query Match 2.1%; Score 155.5; DB 2; Length 1199;
Best Local Similarity 19.4%; Pred. No. 0.12;
Matches 222; Conservative 141; Mismatches 361; Indels 421; Gaps 57;

QY 63 E Q R A W D L A H T W E M G L R S L C A Q A O E G A G H S P S F - - - - - P Y S P S E P H L G S P S Q P T S T A V I M 118
Db 43 D E S S W K - - - H F W - - - - - S L C V N V A A A A F T K F R R Q O D N K R Y K S A - L S L S P P T S V S R I - 91

Fri Jan 30 10:53:59 2004

us-09-996-617-2.rpt

Page 10

Db 918 SPPBI 922

Search completed: January 29, 2004, 13:49:45
Job time : 40.197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 21.9982 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWRLACTYLEFKKE.....HLIMELWKKSKKLLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7502	99.6	1473	1	NAL1_HUMAN
2	1382.5	18.4	1062	1	PS9046 mus sapien
3	1279	17.0	1034	1	CTS1_HUMAN
4	1254.5	16.7	1033	1	CISI_MOUSE
5	1037	13.8	994	1	NAL4_HUMAN
6	983	13.0	1200	1	MATE_HUMAN
7	948	12.6	1062	1	NAL2_HUMAN
8	908.5	12.1	980	1	PYA3_HUMAN
9	901.5	12.0	892	1	PYA5_HUMAN
10	895	11.9	854	1	PYA5_RAT
11	882	11.7	1111	1	MATE_MOUSE
12	868	11.5	843	1	PYA6_MOUSE
13	711	9.4	1033	1	PYA6_HUMAN
14	668	8.9	431	1	CARF_HUMAN
15	393.5	5.2	1040	1	CARF_HUMAN
16	392	5.2	1020	1	CARF_MOUSE
17	388.5	5.2	953	1	CAR4_HUMAN
18	361.5	4.8	456	1	RINI_RAT
19	358	4.8	460	1	RINI_HUMAN
20	354.5	4.7	456	1	RINI_PIG
21	351	4.7	953	1	CAR4_MOUSE
22	345	4.6	1130	1	C27A_HUMAN
23	318	4.2	1155	1	C27A_MOUSE
24	253.5	3.4	193	1	ASC_MOUSE
25	250.5	3.3	195	1	ASC_HUMAN
26	215	2.9	1024	1	CARG_HUMAN
27	212.5	2.8	1403	1	BIRE_HUMAN
28	200	2.7	1403	1	BIRE_MOUSE
29	199	2.6	1402	1	BIRG_MOUSE
30	198	2.6	1403	1	BIR1_HUMAN
31	196.5	2.6	1447	1	BIRB_MOUSE
32	196	2.6	1403	1	BIRA_MOUSE
33	161	2.1	483	1	YAK2_HUMAN

34	158	2.1	5065	1	EPPL_HUMAN
35	152	2.0	3511	1	MY15_MOUSE
36	147.5	2.0	793	1	STSA_MOUSE
37	146.5	1.9	1524	1	Y133_HUMAN
38	145.5	1.9	793	1	STSA_RAT
39	144	1.9	2300	1	CYAA_NEUCR
40	142.5	1.9	4303	1	PKD1_HUMAN
41	142	1.9	1426	1	NPH4_HUMAN
42	142	1.9	1839	1	CYAA_SACKL
43	141.5	1.9	1021	1	PSKR_DAUCA
44	140	1.9	794	1	STSA_BOVIN
45	138.5	1.8	794	1	STSA_HUMAN

ALIGNMENTS

RESULT 1

NAL1_HUMAN STANDARD; PRT; 1473 AA.

AC Q9C000; Q9BZ28; Q9BZ29; Q9HAY6; Q9UFT4; Q9YZE0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NACHT, LRR- and PYD-containing protein 2 (Death effector filament-forming ced-4-like apoptosis protein) (Nucleotide-binding domain and caspase recruitment domain) (Caspase recruitment domain protein 7).
GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIA0926.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21169419; PubMed=11270363;
RA Bertin J., Distefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation proteins";
RL Cell Death Differ. 7:1273-1274 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21148093; PubMed=11250163;
RA Martinon F., Hofmann K., Tschopp J.;
RT "The PYRIN domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation";
RL Curr. Biol. 11:R118-R120 (2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISUB=Erythrocytosis;
RA MEDLINE=21153743; PubMed=11076957;
RA Hlaing T., Guo R.-F., Dilley K.A., Louisa J.M., Morrish T.A., Shi M.M., Vincenz C., Ward P.A.;
RT "Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis proteins";
RL J. Biol. Chem. 276:9230-9238 (2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1, 2; 3 AND 4), AND PROTEIN INTERACTION. TISUB=T-cell;
RX MEDLINE=21153744; PubMed=11113115;
RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S., Godzik A., Reed J.C.;
RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome c-dependent caspase activation and apoptosis";
RL J. Biol. Chem. 276:9239-9245 (2001).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2). TISUB=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE=Uterus;
 RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases
 CC -1- FUNCTION: Able to form cytoplasmic structures termed death
 effector filaments. Enhances APAF1 and cytochrome c-dependent
 activation of pro-caspase-9 and consecutive apoptosis. Seems to
 bind ATP.
 CC -1- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 and with APAF1 in a cytochrome c-inducible way leading to the
 formation of an apoptosome. This interaction may be ATP-dependent.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=NAC beta, DEFCAP-L;
 CC IsoId=Q9C000-1; Sequence=Displayed;
 CC Name=2; Synonyms=NAC alpha, DEFCAP-S;
 CC IsoId=Q9C000-2; Sequence=VSP_004327;
 CC Name=3; Synonyms=NAC gamma;
 CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
 CC Name=4; Synonyms=NAC delta;
 CC IsoId=Q9C000-4; Sequence=VSP_004326;
 CC -1- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 expressed in peripheral blood leukocytes, chronic myelogenous
 leukemia cell line K-562, followed by thymus, spleen and heart.
 CC Also detected in lung, placenta, small intestine, colon, kidney,
 liver and muscle.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 1 CARD domain.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF298548; AAG15254.1; -;
 DR EMBL: AF310105; AAG30288.1; -;
 DR EMBL: AF229059; AAK00748.1; -;
 DR EMBL: AF229060; AAK00749.1; -;
 DR EMBL: AF229061; AAK00750.1; -;
 DR EMBL: AF229062; AAK00751.1; -;
 DR EMBL: AB023143; BAA76770.1; -;
 DR EMBL: AL117470; CAB55945.1; -;
 DR PIR: T1255; T1255.
 DR HSSP: P13489; 1A4Y.
 DR MIM: 606636; -;
 DR GO: GO:0005622; C:intracellular; IC.
 DR GO: GO:0016506; F:apoptosis activator activity; NAS.
 DR GO: GO:0008656; F:caspace activator activity; NAS.
 DR GO: GO:0019899; F:enzyme binding activity; IPT.
 DR GO: GO:0006919; P:caspace activation; NAS.
 DR GO: GO:0006917; P:induction of apoptosis; NAS.
 DR InterPro: IPR001315; CARD
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007091; LRR_RNinh.
 DR InterPro: IPR007111; NACHT_NTPase.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF00560; LRR; 2.
 DR Pfam: PF02758; PAAD_DAPIN_1.
 DR PRINTS: PRO0364; DISEASERESIST.
 DR PROSITE: PS50209; CARD; 1.
 DR PROSITE: PS50824; DAPIN; 1.
 DR PROSITE: PS50837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;

KW Alternative splicing.
 FT DOMAIN 1 92
 FT REPEAT 328 637
 FT REPEAT 704 725
 FT REPEAT 807 830
 FT REPEAT 864 887
 FT REPEAT 921 944
 FT REPEAT 950 973
 FT REPEAT 1199 1215
 FT REPEAT 1216 1236
 FT DOMAIN 1374 1463
 FT NP BIND 314 341
 FT VARSPIC 958 987
 FT VARSPIC 1262 1305
 FT MUTAGEN 340 340
 FT MUTAGEN 340 340
 FT CONFLICT 155 155
 FT CONFLICT 246 246
 FT CONFLICT 782 782
 FT CONFLICT 878 878
 FT CONFLICT 995 995
 FT CONFLICT 1119 1119
 FT CONFLICT 1184 1184
 FT CONFLICT 1241 1241
 FT CONFLICT 1366 1366
 SQ SEQUENCE 1473 AA; 165865 MW; 438FDCE45C2562D CRC64;
 Query Match 99.6%; Score 7502; DB 1; Length 1473;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAGAAGRLACYLEFKEELKEFOLLANKHSSSSGETPAQPEKTSMEVASYLVQA 60
 DB 1 MAGAAGRLACYLEFKEELKEFOLLANKHSSSSGETPAQPEKTSMEVASYLVQA 60
 QY 61 YGBORAMDLALTWEOMGLRSLCAQOEGAGHSPSPYSPSEPHLGSPOPTSTAVLMEV 120
 DB 61 YGBORAMDLALTWEOMGLRSLCAQOEGAGHSPSPYSPSEPHLGSPOPTSTAVLMEV 120
 QY 121 IHELPAGCTQGSERRVRLQPTSGRRWRISASLLYOLPSPDHESPOSPNPTST 180
 DB 121 IHELPAGCTQGSERRVRLQPTSGRRWRISASLLYOLPSPDHESPOSPNPTST 180
 QY 181 AVLGSGWSPPOPSLAPEQEAEGTQWPLETSGIYITEIREREREKSGRPMVAVGT 240
 DB 181 AVLGSGWSPPOPSLAPEQEAEGTQWPLETSGIYITEIREREREKSGRPMVAVGT 240
 QY 241 PPOAHTSLQPHHPWEPVSRESLCTWPKNEDFNOKFTOLLILQRPFRSQDPLVRSW 300
 DB 241 PPOAHTSLQPHHPWEPVSRESLCTWPKNEDFNOKFTOLLILQRPFRSQDPLVRSW 300
 QY 301 PPIVEENRGLLEIRDLFGGLDIOEPRIIVIIQGAAGIKSTLAAVVKAMRGOLYGR 360
 DB 301 PPIVEENRGLLEIRDLFGGLDIOEPRIIVIIQGAAGIKSTLAAVVKAMRGOLYGR 360
 QY 361 FOHVFFYSRELAQSKVSLAELIGKDGATPAPIROLISREERLFLIDGVDEPMVLQ 420
 DB 361 FOHVFFYSRELAQSKVSLAELIGKDGATPAPIROLISREERLFLIDGVDEPMVLQ 420
 QY 421 EESSSELCLHWSPOPADALGSLGKTIIPASFLITARTTALONLIPSEDAWVEVIG 480
 DB 421 EESSSELCLHWSPOPADALGSLGKTIIPASFLITARTTALONLIPSEDAWVEVIG 480
 QY 481 FESSRKEYFYRYFDERQAIAPAFRLVSKNEMALCLPVWSMLACTCLMOOMRKEKL 540
 DB 481 FESSRKEYFYRYFDERQAIAPAFRLVSKNEMALCLPVWSMLACTCLMOOMRKEKL 540
 QY 541 TLTSKTTTTLCHAYLAQALQAPLQPLQRLDCLSLAEGIWOKKTLFSPDDLKHGIDGAI 600
 DB 541 TLTSKTTTTLCHAYLAQALQAPLQPLQRLDCLSLAEGIWOKKTLFSPDDLKHGIDGAI 600

QY 601 ISTEKMGILIOEHPIPIPSYFIHLCOFOFFAAMSIVLEDEKRGKHSNCITIDEKTEAY 660
 DB 601 ISTEKMGILIOEHPIPIPSYFIHLCOFOFFAAMSIVLEDEKRGKHSNCITIDEKTEAY 660
 QY 661 GINGHGFASATSTRFLGLISDEGEREMENIFHCRLSOGNIMQWVPSIQLLQPHSLSLH 720
 DB 661 GINGHGFASATSTRFLGLISDEGEREMENIFHCRLSOGNIMQWVPSIQLLQPHSLSLH 720
 QY 721 CLYTRKKTFLTOVMAHEEWKCEVETDMELLVCTPCKFSRHVYKQQLIEGRORSTWS 780
 DB 721 CLYTRKKTFLTOVMAHEEWKCEVETDMELLVCTPCKFSRHVYKQQLIEGRORSTWS 780
 QY 781 PTMVVLFRWVPTDPAVMOILFSYVKATRNKELDLGNSLSHSAVKSLCKTLRPPCLTE 840
 DB 781 PTMVVLFRWVPTDPAVMOILFSYVKATRNKELDLGNSLSHSAVKSLCKTLRPPCLTE 840
 QY 841 TLRLAGCGLTAECDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCORLPOPSCKLQRLQ 900
 DB 841 TLRLAGCGLTAECDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCORLPOPSCKLQRLQ 900
 QY 901 LVSGGLTSDCCODLAVLSASPSLKEHDLQNNITDVGVRLLCGLRHPACKLIRLGLDQ 960
 DB 901 LVSGGLTSDCCODLAVLSASPSLKEHDLQNNITDVGVRLLCGLRHPACKLIRLGLDQ 960
 QY 961 TTLSDEKROELRAEOEKPOLLIIFRRRPSVMTPEGLDTGEMSNSTSLKROBLGSSRA 1020
 DB 961 TTLSDEKROELRAEOEKPOLLIIFRRRPSVMTPEGLDTGEMSNSTSLKROBLGSSRA 1020
 QY 1021 ASHVAQANLKLIDVSKIFPIAIEAESSEPEVVPYELLCPSPASQGLDHTKPLGTDDEFW 1080
 DB 1021 ASHVAQANLKLIDVSKIFPIAIEAESSEPEVVPYELLCPSPASQGLDHTKPLGTDDEFW 1080
 QY 1081 GPTGPVATEVVDKKNLYRHVFPVAGSYRMVNTGLCFVMEATVTEIEFCVWDQFLGBIN 1140
 DB 1081 GPTGPVATEVVDKKNLYRHVFPVAGSYRMVNTGLCFVMEATVTEIEFCVWDQFLGBIN 1140
 QY 1141 POSHMMVAGPLDIDKABGAYEAHLHPFVALQGHVDTSLFQMAHFKESGMLLEXPARY 1200
 DB 1141 POSHMMVAGPLDIDKABGAYEAHLHPFVALQGHVDTSLFQMAHFKESGMLLEXPARY 1200
 QY 1201 ELHHIVLENSPSPGLVLLKMHNAALRFIPVTSVLLYHRVHPREVTFLYLIPSDCSIR 1260
 DB 1201 ELHHIVLENSPSPGLVLLKMHNAALRFIPVTSVLLYHRVHPREVTFLYLIPSDCSIR 1260
 QY 1261 -----KELELCYRSFGEDQLF 1276
 DB 1261 KAIDLEMKQFVRIHPRPLTPLYMGCRVTVSGSGSMLEILKRELCLCRSPGEDQLF 1320
 QY 1277 SEFVVGHLGSGIRLOVQDKDETLWEALVYKPGMLPATTLIPARIAVPSPLDAPQLH 1336
 DB 1321 SEFVVGHLGSGIRLOVQDKDETLWEALVYKPGMLPATTLIPARIAVPSPLDAPQLH 1380
 QY 1337 FVDDYRQQLARVTSVEVLDKLGQVLSQBOYRYLAENTRPSQMKLFSLSQSMRKC 1396
 DB 1381 FVDDYRQQLARVTSVEVLDKLGQVLSQBOYRYLAENTRPSQMKLFSLSQSMRKC 1440
 QY 1397 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1429
 DB 1441 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1473
 RESULT 2
 PYAT_HUMAN STANDARD; PRT; 1062 AA.
 AC P59046;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PPRIN-containing APAF1-like protein 7 (Monarch-1).
 GN PYPAF7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Meriam S.,
 RA Iora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.,
 RT "PYPAF7, a novel PPRIN-containing APAF1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing.";
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC TISU8-Jymphoma;
 RA Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISU8-Leukocyte;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield J.Y.S.N., Krzywinski M.I., Skalske J., Smillie D.E.,
 RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 CC activation of NF-kappa-B via IKK.
 CC -1- SUBUNIT: Binds to ASC with its DAPIN domain.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=1;
 CC IsoId=P59046-1; Sequence=Displayed;
 CC Name=2; Synonyms=11;
 CC IsoId=P59046-2; Sequence=VSP_005524;
 CC Name=3; Synonyms=111;
 CC IsoId=P59046-3; Sequence=VSP_005523;
 CC -1- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
 CC predominantly in eosinophils and granulocytes, and at lower levels
 CC in monocytes.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AY095146; AAM18327.1; -
 CC EMBL: AY116204; AAM75142.1; -
 CC EMBL: AY116205; AAM75143.1; -
 CC EMBL: AY116206; AAM75144.1; -
 CC EMBL: BC028069; AAM75144.1; -
 CC EMBL: BC028069; AAM75144.1; -
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR007091; LRR_RNinh.

phenotype including recurrent fever, cold sensitivity, sensorineural deafness, and AA amyloidosis.";

Arthritis Rheum. 46:2445-2452(2002).

[4]

SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).

TSSUSE=BI0d;

MDLINE=2049367; PubMed=11042152;

Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G., Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W., Rao Y., Huang Q.-H., Zhou Y., Hu G.-X., Chen S.-J., Chen Z.;

"Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells.";

Genome Res. 10:1546-1560(2000).

[5]

VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND ARG-569, AND VARIANT FCAS/MWS TRP-260.

MDLINE=21987640; PubMed=11922556;

Dode C., Le Du N., Cutsbet L., Lecomte F., Berthelot J.-M., Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A., Granel B., Frances C., Garcier F., Ebery P., Boulanger S., Dougnes J.-P., Delpech M., Gatteau G.;

"New mutations of CIAS1 that are responsible for Muckle-Wells syndrome and familial cold urticaria: a novel mutation underlies both syndromes.";

Am. J. Hum. Genet. 70:1498-1506(2002).

[6]

VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND THR-662, AND TSSUSE SPECIFICITY.

MDLINE=22062556; PubMed=12032915;

Feldman J., Ptiere A.-M., Quantier P., Berguin P., Certain S., Corliss E., Teillac-Hamel D., Fischer A., de Saint Basile G.;

"Chronic infantile neurological cutaneous and articular syndrome is caused by mutations in CIAS1, a gene highly expressed in polymorphonuclear cells and chondrocytes.";

Am. J. Hum. Genet. 71:198-203(2002).

-1- FUNCTION: May function as a potential inducer of apoptosis. Interacts selectively with apoptosis-associated specklike protein containing a CARD domain (ASC). This complex may function as an upstream activator of NF-kappaB signaling.

-1- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=3;

Name=2;

isoId=Q96P20-1; Sequence=Displayed;

Name=1;

isoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;

Name=3;

isoId=Q96P20-3; Sequence=VSP_005519;

-1- TISSUE SPECIFICITY: Expressed in blood, leukocytes. Strongly expressed in polymorphonuclear cells, undetectable or expressed at a lower magnitude in B and T lymphoblasts, respectively. High level of expression detected in chondrocytes. Low or no expression in the other tissues tested.

-1- DISEAS: Defects in CIAS1 are a cause of familial cold autoinflammatory syndrome (FCAS), commonly known as familial cold urticaria. FCAS is rare autosomal dominant systemic inflammatory disease characterized by episodes of rash, arthralgia, fever and conjunctivitis after generalized exposure to cold.

-1- DISEASB: Defects in CIAS1 are a cause of Muckle-Wells syndrome (MWS), a rare autosomal dominant fever syndrome with episodic urticaria, arthralgia, amyloidosis and progressive sensorineural deafness.

-1- DISEAS: Defects in CIAS1 are the cause of chronic infantile neurologic cutaneous and articular syndrome (CINCA), also known as neonatal onset multisystem inflammatory disease, or NOMID, a rare congenital inflammatory disorder characterized by a triad of neonatal onset of cutaneous symptoms, chronic meningitis, and joint manifestations with recurrent fever and inflammation.

-1- SIMILARITY: Contains 1 NACHT domain.

-1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

-1- CAUTION: Ref.4 sequence differs from that shown due to frameshifts in positions 893, 918 and 926.

CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL; AF410477; AAL33908.1; -;	
DR	EMBL; AF427617; AAL33911.1; -;	
DR	EMBL; AY051117; AAL12487.1; JOINED.	
DR	EMBL; AY051112; AAL12487.1; JOINED.	
DR	EMBL; AY051113; AAL12497.1; JOINED.	
DR	EMBL; AY051114; AAL12497.1; JOINED.	
DR	EMBL; AY051115; AAL12497.1; JOINED.	
DR	EMBL; AY051116; AAL12497.1; JOINED.	
DR	EMBL; AY056059; AAL12487.1; JOINED.	
DR	EMBL; AY056060; AAL12487.1; JOINED.	
DR	EMBL; AY051117; AAL12498.1; -;	
DR	EMBL; AY051112; AAL12498.1; JOINED.	
DR	EMBL; AY051113; AAL12498.1; JOINED.	
DR	EMBL; AY051114; AAL12498.1; JOINED.	
DR	EMBL; AY051115; AAL12498.1; JOINED.	
DR	EMBL; AF420469; AAL65136.1; -;	
DR	EMBL; AF420469; AAL65136.1; ALT INIT.	
DR	EMBL; AY092033; AAM14669.1; ALT INIT.	
DR	EMBL; AF418985; AAL14640.2; ALT INIT.	
DR	EMBL; AF054176; AAC39910.1; ALT_FRAME.	
DR	Genew; HGNC:16400; CIAS1.	
DR	MIM; 606416; -;	
DR	MIM; 120100; -;	
DR	MIM; 120100; -;	
DR	MIM; 121900; -;	
DR	MIM; 607315; -;	
DR	GO; GO:0016306; F:apoptosis activator activity; NAS.	
DR	GO; GO:0006917; P:induction of apoptosis; NAS.	
DR	GO; GO:0006954; P:inflammatory response; IMP.	
DR	GO; GO:0007165; P:signal transduction; NAS.	
DR	InterPro; IPR001611; LRR.	
DR	InterPro; IPR007091; LRR_RNinh.	
DR	InterPro; IPR003590; LRR_RNinh_sub.	
DR	InterPro; IPR007111; NACHT_NTPase.	
DR	InterPro; IPR004020; PAAD_DAPIN_dom.	
DR	Pfam; PF00560; LRR; 2.	
DR	SMART; SMO0368; LRR_RI_3.	
DR	PROSITE; PS50824; DAPIN; 1.	
DR	PROSITE; PS50837; NACHT; 1.	
KV	Apoptosis; Repeat; Leucine-rich repeat; Alternative splicing; Disease mutation; Deadness.	
KW	DOMAIN 1 91 DAPIN.	
FT	DOMAIN 218 534 NACHT.	
FT	REPEAT 738 761 LRR 1.	
FT	REPEAT 795 818 LRR 2.	
FT	REPEAT 852 875 LRR 3.	
FT	REPEAT 881 904 LRR 4.	
FT	REPEAT 909 937 LRR 5.	
FT	REPEAT 939 961 LRR 6.	
FT	REPEAT 966 989 LRR 7.	
FT	DOMAIN 688 695 POLY-GLU.	
FT	VARSPLIC 718 1034 Missing (in isoform 3).	
FT	VARSPLIC 719 775 /Frtid=VSP_005519. Missing (in isoform 1).	
FT	VARSPLIC 834 890 /Frtid=VSP_005520. Missing (in isoform 1).	
FT	VARSPLIC 834 890 /Frtid=VSP_005521. Missing (in isoform 1).	
FT	VARIANT 198 198 V -> M (IN FCAS AND MWS).	
FT	VARIANT 260 260 R -> W (IN FCAS AND MWS).	
FT	VARIANT 303 303 /Frtid=VAR_014104 D -> N (IN CINCA AND MWS).	
FT	VARIANT 305 305 /Frtid=VAR_014105. L -> P (IN FCAS AND MWS).	

FT REPEAT 908 931 LRR 5.
 FT REPEAT 937 964 LRR 6.
 FT REPEAT 965 988 LRR 7.
 SQ SEQUENCE 1033 AA; 118274 MW; 5924690966812117 CRC64;

Query Match 16.7%; Score 1254.5; DB 1; Length 1033;
 Best Local Similarity 32.6%; Pred. No. 1,2e-72;
 Matches 331; Conservative 164; Mismatches 334; Indels 187; Gaps 23;

QY 8 RLACTLEFLKEELKEFOLLANKAHSSSSGSETAQOPEKTSQGVAVSYLAQVGEQAM 67
 DB 7 KLAQLELELEVDLKKFKHLEDEYPEREKGCIPVPRQMEKMDHDLATLMDFNEEKAM 66
 QY 68 DLATHTWQMGLRSACAOBAGHSPSPYSPSPHSGSPQSTAVMPWHEHLAG 127
 DB 67 AMAVIFALINRDMEKAKK-----DQ-----EANDT 95
 QY 128 CTGSGRRVRLQLPDTSGRRRREISASLLYQALPSSPDHESPQSPNAPTSTAVLGSWG 187
 DB 96 CTSHS-WVCGE--DSLEEW-----MGLLG 118
 QY 188 SPPOSLAPROEAGTQWPLDTSIGIYTEIREEREKSEKRRPMAAVGTPOANTS 247
 DB 119 YLSRISICKKKK-----DYCKMYRRHVRSFYSIKQRM-----ARLG----- 155
 QY 248 LQPHHNPSPVRESLSTWPKNEDFNQKFTQLLLQRPHPRSQDPLVKSMPDYVEN 307
 DB 156 -----ESVDLSRYTQLQV-KENHSKOE----- 178
 QY 308 RGHLL-----EIRD-----LFGGLDTPQRP-RIVILQGAIGKSTLAEQVXA 350
 DB 179 REHELITIGRTWNRDSPMSLKLLEFEPEDGHSFVHTVYVQGAAGIKTILANKIMLD 238
 QY 351 WGRGQYGDGRPOHVVYFSCRELQASKVVSALIGKDGATAPAPRQILSRPELLFLD 410
 DB 239 WALGTLFKDKDYLFPIHCREVSLRTPRSLADLIVSCMPDPNPVCKILRKRSRILFLMD 298
 QY 411 GVDERGWLQOEPSSSLCHWSQOPADALIGSLKTLTPKASPLTARTTLOQLISL 470
 DB 299 GFDELQGAFFDEHIGVCTDMQKAVNGDILLSSILKKLIPKASLLITTRPVALEKQHL 358
 QY 471 EQARWVAVLGFSSSRKEFYRYFTDERQATRAFLVYSNKLMLALCLVPWVSACTCL 530
 DB 359 DHPRAVEILGFSEAKRKEFYFKEFVSNELQARAFILQENVELFTMCFPLVCWVCGGL 418
 QY 531 MQQMKRKEKTLTSTTTTLCIHYLAQALQAP-----LGPQLRDLCSLAAGIWKQKT 584
 DB 419 KQOETGKSLAQTSKTTTAVVVFLLSSILQSRGIEEHLFSDYLOGLCSLADGIWQKI 478
 QY 585 LPSPDLLKRGHLDGAILISTELKMGILQ-EHPIPLSYSPHLCOFEPFAMSVLEDE--- 640
 DB 479 LFEBCDLKRGHGLQKTVDVSAFLRMNVFOKEVDCERFYSFSHMTQSFAMTYLLLEEBAG 538
 QY 641 ----KGRGKHSNCI-IDLEKTLVAVG-INGLFGASTTFFLGLLDESGREMEINIFHRL 694
 DB 539 ETVKRGPGGCGDLNADVKNVLENGKFEKGLIIVAFELVAVOERTSYLEKLSCKI 598
 QY 695 SQ--GRNIMQWV---PSLQILLQPHSLBSLCLVETRNKTFPLTVMAHFEEMGCVETD 748
 DB 599 SQOQVLELKLTEVAKKAKKQWQPSQLELFCVLEMOBEDFVQAMDFPRTIEINLSTR 658
 QY 749 MELVCTCICKSRNVKQLQ-----IEGRHSTMSPT-----MVLFPWV 790
 DB 659 MDHVVSSFLKNCGRVKTSLGFFPNSPEKEEBEERGGRPLOVOCVPRPDTHVACSSRLV 718
 QY 791 P--VTDAWQILFVYLKTVRNLEKELSGNSLSHSAVSKLCTTLRPRCLLETLAAGC 848
 DB 729 NCCLLTSSFCRGFSSLSSTRSLTELDLSONTLGDGPMVTLGALHPGNCILRLMLGRCG 778
 QY 849 LTAEBCKDLAFGLRANQTLTELDLSFNVLTDAAGHLCORLQFQSCKQRLQVSCGLTS 908
 DB 779 LSHQCCFIDISSVLSQKLVLELDLSDNALGDGIRBLCVGLKHLNCLQKMLVSCCLTS 838

QY 909 DCCODLASVSPSLKELDLQNNLDVGYRLCEGHRPACKLIRGLDPTTIS 964
 DB 839 ACCODLALVLSNHSNLSRLRYIGNALBDSGVQVLCERKMDPQCMLQKGLVNSGLT 894

RESULT 5

NAL4_HUMAN
 ID_NAL4_HUMAN STANDARD; PRT; 994 AA.

AC Q96NM2; Q96AY6; 41, Created
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NACHT, LRR- and PYD-containing protein 4 (PAD and NACHT-containing protein 2) (PYRIN-containing APAF1-like protein 4) (ribonuclease inhibitor 2).
 GN NALP4 OR PYPA4 OR PAN2 OR RNH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1] SEQUENCE FROM N.A. (ISOFORM 1).
 RA Martinon F., Tschopp J.,
 RT "NALP4, a novel member of the PYD, NACHT, and LRR family,"
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RN [2] SEQUENCE FROM N.A. (ISOFORM 1).
 RA Florentino L., Reed J.C.,
 RT "Pan2, a novel PAD-containing protein,"
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 RN [3] SEQUENCE FROM N.A. (ISOFORM 1).
 RA Bertin J.,
 RT "PYPA4, a novel PYRIN-containing APAF1-like protein,"
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RN [4] SEQUENCE FROM N.A. (ISOFORM 1).
 RA Miyamoto T.,
 RT Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RN [5] SEQUENCE FROM N.A. (ISOFORM 3).
 RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsura N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto U., Isono Y.,
 RA Katsuo K., Nakamura Y., Sekine M., Kikuchi H., Yamashita H.,
 RA Kawanishi K., Takahashi-Fujii A., Oshima K., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.,
 RT "NEBO human cDNA sequencing project,"
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN [6] SEQUENCE OF 436-994 FROM N.A. (ISOFORM 2).
 RP TISSUE=Placenta;
 RC MEDLINE=22388257; PubMed=12477932;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Scheet C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopfner R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrion P., Prange C.J.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
 RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length

Db 954 VLGRKTDPEDETOALLTAEEERNPNLTI-----TDODDT 988

RESULT 6

MATE HUMAN STANDARD; PRT; 1200 AA.

AC P59047;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mater protein homolog.
GN MATER.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RL Hum. Reprod. 17:903-911(2002).
CC -1- TISSUE SPECIFICITY: Oocyte-specific.
CC -1- SIMILARITY: Contains 1 DAPIIN domain.
CC -1- SIMILARITY: Contains 1 NACTH domain.
CC -1- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC -1- CAUTION: It is not obvious that this is the ortholog of mouse
Mater.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC -----
DR EMBL; AY054966; AAL1549.1; -
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACTH_NTPase.
DR InterPro; IPR004020; PAAD_DAPIIN_dom.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF02758; PAAD_DAPIIN; 1.
DR PRINTS; PRO0019; LEUCRICHPT.
DR SMART; SMO0368; LRR_R1; 11.
DR PROSITE; PS50824; DAPIIN; 1.
DR PROSITE; PS50837; NACTH; 1.
DR ATP-binding; Leucine-rich repeat; Repeat.

KW DOMAIN 57 148 DAPIIN.
FT REPEAT 280 602 NACTH.
FT REPEAT 704 727 LRR 1.
FT REPEAT 730 753 LRR 2.
FT REPEAT 780 803 LRR 3.
FT REPEAT 809 832 LRR 4.
FT REPEAT 836 863 LRR 5.
FT REPEAT 865 892 LRR 6.
FT REPEAT 893 916 LRR 7.
FT REPEAT 950 973 LRR 8.
FT REPEAT 979 1002 LRR 9.
FT REPEAT 1007 1034 LRR 10.
FT REPEAT 1036 1059 LRR 11.
FT REPEAT 1064 1092 LRR 12.
FT REPEAT 1121 1142 LRR 13.
FT NP_BIND 286 293 ATP (POTENTIAL).
SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A771B28FA CRC64;

Query Match 13.0%; Score 983; DB 1; Length 1200;
Best Local Similarity 25.4%; Pred. No. 4,3e-55;

Matches 329; Conservative 176; Mismatches 411; Indels 380; Gaps 39;

QY 5 AMGRACYLEPLKEBEIKEROLLANKARRSSSGETPAPKEKTSGEVASYVAQ-YGE 63
DB 62 SYGLQWCLYE-LKEEFQTFKELKKKS-SESTTCSIPEFIEIANVECLALLHERYGA 119
QY 64 QRAMDLALHTWEQGLRSLCAQOEGAGSPSPYSSEPHLSGPSQPTSTAVLMPWTHE 123
DB 120 SLMAWTSISIFEMNNLTLEKARDMKR-----HSDEDEATMTDQPS-----KE 166
QY 124 LPAGCTGSSRRVRLQPLDTSGRWRREISASLLYQALPSSPDHSPQESPNAPTAVL 183
DB 167 KVPGISQAVQQ-----DSATAETKQGISQA----- 193
QY 184 GSNQSPQPSLARREQAPGTQWPLDTSGLTYTEIRERRESEKRRPMAAVGTPPQ 243
DB 194 -----MEQGA-----TAATEBOEISQAEQGAATAETEEQ 226
QY 244 AHTSLQPHHPWEPVSVESLSTPWNEDFNOKFTQLLLORPHRPSODPLVRS----- 299
DB 227 GHGG-----DTWDYKSHWT-KF-----AEEEDVRSFENT 256
QY 300 ---WPDVEENRGHLIEIRDLPFGGLDTPQ---EPRVILQGAAGIGKSTLARGKEAWG 352
DB 257 AADWP-----EMQTLAG-AFSDRWGRFRPRVVLHGKSGIKSALARRIVLCWA 304
QY 353 RGQYGDPRFQVHYFYSCRELAQSKVVSALBKGDTATPAPRQLISREPRLLFIIDGV 412
DB 305 QGGIYQGMFVYFVPLPRERQKKESSVTEFIREWDSQAPVTEISRBERLLFIIDGF 364
QY 413 DEPGWVLAQPSSESLCHWSPQPADALGSLGKTIIPESAFITARTTALQNLIPLEQ 472
DB 365 DDGASVLAN-DTKLCKDMARKQPPFTLIRSLKVLPEPFLVTVADVGTETKSEVVS 423
QY 473 ARWVEVLGFSESSRKEYFYFYFDERQAIRAFILVSNKELMALGLVPWWSWLAQCLMQ 532
DB 424 PRYLTVAGISGEORIHLLERIGEHOQTGLBAIINRRLDQCQPAVAGSLICVALQL 483
QY 533 QMRKEKTLTSTKTTTLCIHYLAQALQAPLG-----POLRDLCSLAABEIMQ 581
DB 484 QDVVGESVAPFNQTLTG--LHAAPAFHQLTPRGVRRCLNLEERVVLRKRCMAVEGVN 541
QY 582 KTLFSPDDLRKKGDLDAISTELKMGILIOEHPIPLS-----YSFIHLQFOEPPAASVY 636
DB 542 KSVTFQDDLMVQGLGSEELRALFHMNL-----LPDSHCERYTFFHLSQDFPALYYV 597
QY 637 LEDEKGRGKHSNCIIDEKTLLEA-----YGIHGLFGASTRPFLLGLSDEGERENI 689
DB 598 LEGLE-IBPALCPLYVEKTRSMELKQAGPHIHSILW---MKRFLFGLVSEVDVRRPLEV 652
QY 690 FHCRLSQG--RIMQWVPSLQLLQPH-----SLBSLHCLVETRNKTPILQVAAHFEEM 741
DB 653 LGCPVPVGVQKLIHWSSLLG--QQPNATTPGDTLDAFHLFETQDEKFPRLALANSQEV 710
QY 742 GMCVETMELLYVCFCIKFGRHYVKQLQ-LEG---RQHRSTWSPVTVLFRWVPTDAYV 797
DB 711 WLPINQWDLIASFCLOHCPYIKRIVDVKGIFPRDESAAEACPVVILNMRDKTLIEQW 770
QY 798 -----OLIF 801
DB 771 EDFCSMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNMQTFPGVHLM 830
QY 802 SVLKVTNKLKELDLSGNSLSHSAVSKSLCTKLRBRCLETLRLAGCGLT----- 850
DB 831 RIVANNNRSLNMGTHLKEEDVRVACBALKPKKCLLESIRLDCCGLTHACVYKISQIL 890
QY 851 -----ABD----- 853
DB 891 TTSPSLKSLSLAQNKYTDQGVTPLSDALRVSCALQGLIEDCGITATGCGASLASALVN 950
QY 854 -----CKDLAFLGRANQTL 867
DB 951 RSLTLCLSNNSLGNBGNLLCRSMRLPHCSLQRLMLNQCHLDPTAGCGSIALALMGNSWL 1010

QY 868 TELDLSFNVLTDAAGAKELCORLROPSCKLORLOVSGELTSDCCODLASVLSAPSLKEL 927
 DB 1011 THSLSNPVEDNGVKLLCEWMBPSCHLQDLIELVKCHLTAACCESISCVTSRRLKSL 1070
 QY 928 DLQONNLDVGVKLLCGSLHAPACKLRLGLDQDTLSDEKROELRALEQKPOLLIIFSR 987
 DB 1071 DLTDMALGDDGVAALCGELQKQSVLRLGLKACGLTSDCCALST-----ALSCNRH 1123
 QY 988 KPSVMTPEGLDGENSTSTSLKROBLGSEBAASHVAONLKLVDYK-IFP--IAEIA 1044
 DB 1124 LTLT-----NLVONNNSPKKMMKLSAFAC--PTSNLQITGLMKQYVQIRKLL 1171
 QY 1045 EESSPEVVPVELLCVPSASQGLHTKPLGTDDPFW 1080
 DB 1172 EE-----VQLT-KPRVVIDGSWHS--FDEDPRYM 1197

RESULT 7
 NAL2_HUMAN
 ID NAL2_HUMAN STANDARD; PRT; 1062 AA.
 AC Q9NX02; Q9H6V5; Q9H6G6; Q9HAV9; Q9NMK3;
 DT 16-OCT-2001 (Ref. 40, Created)
 DT 16-OCT-2001 (Ref. 40, Last sequence update)
 DT 13-SEP-2003 (Ref. 42, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 2 (Nucleotide-binding site
 protein 1) (PYRIN-containing APAF1-like protein 2).
 GN NALP2 OR NBS1 OR PYPAF2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., D'Stefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 proteins.";
 RL Cell Death Differ. 7:1273-1274(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martison F., Hofmann K., Technop J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 implicated in apoptosis and inflammation.";
 RL Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22162427; PubMed=12019263;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., D'Stefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing APAF1-like protein that regulates
 activation of NF-kappa B and caspase-1-dependent cytokine
 processing.";
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Colon, and kidney epithelium;
 RA Matanabe K., Kunagai A., Itakura S., Yamazaki M., Taahito H., Ota T.,
 RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isogai T., Sugano S.;
 RT "MEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=lung, and Placenta;
 RX STRAUBERG R.L., FEINGOLD E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.B.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millamy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahney J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC - FUNCTION: May be implicated in apoptosis (By similarity).
 CC - COFACTOR: Binds ATP (By similarity).
 CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC - ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9NX02-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NX02-2; Sequence=VSP_005522;
 CC - SIMILARITY: Contains 1 DAPIN domain.
 CC - SIMILARITY: Contains 1 NACHT domain.
 CC - SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF298547; AAG15253.1; ALT_INIT.
 DR EMBL; AF310106; AAG30289.1; -
 DR EMBL; AF464764; AAL69962.1; -
 DR EMBL; AF464764; AAL69962.1; -
 DR EMBL; AK000517; BAA91223.1; -
 DR EMBL; AK025952; BAB15293.1; -
 DR EMBL; AK00784; BAA91377.1; ALT_INIT.
 DR EMBL; BC003592; AAH03592.1; -
 DR EMBL; BC001039; AAH01039.1; -
 DR HSSP; F10775; ZBNH.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR PROSITE; PSS0837; NACHT; 1.
 DR PROSITE; PSS0837; NACHT; 1.
 DR PROSITE; PSS0837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 94 DAPIN.
 FT REPEAT 207 526 NACHT.
 FT REPEAT 467 491 LRR 1.
 FT REPEAT 622 645 LRR 2.
 FT REPEAT 754 777 LRR 3.
 FT REPEAT 810 832 LRR 4.
 FT REPEAT 839 862 LRR 5.
 FT REPEAT 867 890 LRR 6.
 FT REPEAT 924 947 LRR 7.
 FT REPEAT 981 1005 LRR 8.
 FT REPEAT 1010 1033 LRR 9.
 FT NP_BIND 213 220 ATP (POTENTIAL).
 FT DOMAIN 518 523 POLY-GLU.
 FT VARSPLIC 133 154 Missing (in isoform 2).
 FT FTId=VSP_005522.
 FT M -> V (IN REF. 1).
 FT L -> P (IN REF. 1).
 FT R -> S (IN REF. 4; BAB15293).
 FT R -> K (IN REF. 1).
 FT MISSING (IN REF. 1).
 FT A -> E (IN REF. 5).
 FT SEQUENCE 1062 AA; 120514 MW; 4DDB0F6B3C28C8A7 CRC64;

Query Match 12.6%; Score 948; DB 1; Length 1062;
 Best Local Similarity 27.8%; Pred. No. 6,4e-53;
 Matches 302; Conservative 167; Mismatches 389; Indels 228; Gaps 32;

QY 148 WREISASLLYQALDSSPDHSPQSPQSPNAPTSTAVLGSWGSPPQSLAPRQAPGTQMP 207
 DB 67 WVEWASLQVEFKM-----HRMDLSERAKDEVRBAALSFNKRKPLSLGTRKERP-----P 117
 QY 208 LDEISGIVYEIRERREKSEKGRPPWAAYVGTTPQAHTSLQPH-----HHMPSPVRES 262
 DB 118 LD-----VDEMERFTEAQFTETKGNVCLGKEVFYKGRKPKDKRRCRILTKKRE- 170
 QY 263 LCTWPMKNEFDN---OKFTQLLLQRPHPRSQDPLVRSWPDVVEENRGLIRIDLF 319
 DB 171 MKSMFQDSKEVQWMAERYKMLIFPSNP-----RVLPG 203
 QY 320 PGLDTQSPRIYILGGAAGIGKSTLAROVKEAMRGOLYGRFOHVFYFSCRELAQSKVVS 379
 DB 204 PFSYT-----VVLGPAGLGTTLAQKLMMDWADNLI-HKFKYAFILSCRELSRLGPS 257
 QY 380 LAELIGDGNATPAPRIQOLSRPRLPIFDGVE---PQVTLQESSSELCLHWSQPOP 435
 DB 258 FAELVDFDWELEODDIPIHIAQARKIIFVIDGFDELGAAPGALTE---DICGWEKXKP 313
 QY 436 ADALIGSLGKTLIPASPIITARTALONLPSLEQARWVEVLGFSSSKXEYFYFT 495
 DB 314 VPVLISGLNRWMLPKALLVTTTPRALRDRIIAEETIYIRVGFLEEDRRAVFLHFG 373
 QY 496 DERQAIAPRILVYSKNEKLMALCLVPWVSWLACTLMOQMKREKTLTTSKTTTLCIAYL 555
 DB 374 DEQAMAFELMRNNAALFQGSAPAVCMIVCTTLKQMEKGEDPVTCILTRGLFLRFL 433
 QY 556 AQL- QAPGIPQRLDCSLAEGINOKTILFSPDDIRKGLDCAITSTLXKGILOEHP 614
 DB 434 CSRPQQAQGRGALRTSLAAQQLMAQTSTVHREDERLGVQSSDRRLFDGDIQLQDR 493
 QY 615 IPLS-YSFILCPQEPFAAMSYYL--EDBKRGKSGSCIIDLEKTLBAYGHGFGAST 671
 DB 494 VSKGCVFIFHSFOQFALTALFTLEKEEBEDROCHTDIGDVQKLS--GVERLRNDLI 551
 QY 672 R---FLGLLSDEGEREMENIFHCRLSQGRMLQMWPSL-QLLAQ-----PHSL--- 716
 DB 552 QAGYVSGLANEKRAKLEATFGGRMS-----PDIKELLRCOISCKGSHSTYDL 602
 QY 717 -ESLHCIVERNKFTFLQVNAHEEMGCETDMLLVCFTFCIFSRHAKLQ----- 769
 DB 603 QELGCLYESQEEELVKEVMAQFKEISHLNA--VDVPPSPCVGHKCNLOMSIQVIKEN 661
 QY 770 -----IEGRQ-----HRTSPPTWV----- 784
 DB 662 LPEVNTASBDAVERSDODQHLPTWTDLCSTFGSKNDLMGLAINDSFISASIVRLICE 721
 QY 785 -----VLFRRVPTDVAWQILFSLVKTRNLKELDLSGNSLSHSAVSKLTLR 833
 DB 722 QIASDTGLQRVVKNISPADAHNNCLLA-LRGKTYVLYLQND--QDDMFPLCEVLR 779
 QY 834 RPRCLLTTLRAGGGLTAEDCKDLAFGRANQTTTELDLSNVLTLDGAHCLCRLAQPS 893
 DB 780 HPECNLELYLGLVSCSATYQWADLSLAEVQSLCVLSDNEILLDGAALKLYTLRHPX 839
 QY 894 CKLQRLVSGGLTSDCCODLAVLSASPSLKEILDQNNLDVGVLLCGEGLHAPCKL 953
 DB 840 CFLRLBLENHCLBLANCKDLAAVAVVSRELTTHCLANP;IGNGVFLCGLGLYPECKL 899
 QY 954 IRLGLDQTLSDENROELRALEQEKPOLIFSRKRPVMTPTBGLDGENSNSTSSLRQ 1013
 DB 900 QTLVLMNDITSDCCDLTKLQEKSSLLC-----LDG----- 933
 QY 1014 RLGSERASHVAQANKLIDVSKFPIAET-----AESSPVVEVELLCVPSPASQ--- 1066
 DB 934 -----LNHIGVGMKFLCEALRKPLCNLRCLMLWGCSIPPSCEDLCALSCNQLVLT 986

QY 1067 -DLATKPLGTD-----DDFWGPTGPVATEVVDKENLY---R 1099
 DB 987 LDIQGNPILSSGVKMLFETITCSTSSGTLRTLRKIDDFNDELNKLLEIEKNPQLIIDTE 1046
 QY 1100 VHPVVA 1105
 DB 1047 KHPVVA 1052

RESULT 8

PIA3_HUMAN STANDARD; PRT; 980 AA.

AC O8WX94;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE PYRIN-containing APAF1-like protein 3.

GN PYPAF3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP MEDLINE=22162427; PubMed=12019269;

RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,

RA Lora J.M., Geddes B.U., Brielkin M., DiStefano P.S., Bertin J.,

RT "PIPAF7, a novel PYRIN-containing APAF1-like protein that regulates

RT activation of NF-kappa B and caspase-1-dependent cytokine

RT processing";

RL J. Biol. Chem. 277:29874-29880(2002).

CC -1- SIMILARITY: Contains 1 DAPIN domain.

CC -1- SIMILARITY: Contains 1 NACHT domain.

CC -1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by, and for commercial

CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC

DR EMBL; AF464765; AAL69963.1; -

DR InterPro; IPR007091; LRR_RNinh.

DR InterPro; IPR007111; NACHT_NTPase.

DR InterPro; IPR004020; PAAD_DAPIN_dom.

DR Pfam; PF02758; PAAD_DAPIN; 1.

DR PROSITE; PS50824; DAPIN; 1.

DR PROSITE; PS50837; NACHT; 1.

KW ATP-binding; Leucine-rich repeat; Repeat.

FT DOMAIN 1 93

FT 172 491

FT REPEAT 614 638

FT REPEAT 674 697

FT REPEAT 760 784

FT REPEAT 788 810

FT REPEAT 817 840

FT REPEAT 845 868

FT REPEAT 874 897

FT REPEAT 902 928

FT REPEAT 933 957

FT NP_BIND 178 185

SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;

Query Match 12.1%; Score 908.5; DB 1; Length 980;
 Best Local Similarity 30.2%; Pred. No. 2e-50;
 Matches 261; Conservative 128; Mismatches 284; Indels 191; Gaps 21;

QY 294 PLVKSMPDYVEN-----RGLIIRLD-----LFGGLDTQGER-----IVIL 332
 DB 117 FGEKEGRMSMERQSLWKNQTFWQGDIDNFHDVTLRNQRFIFLNPRTKLTPTTVL 176

```

QY 333 OGAAGIGKSTLAROYKAMGRGOLYGDHFQVFPSCRELAQSKVSLAELIGKDGATP 392
Db 177 HGAAGVGTAKKACMLDWDTCNL-SFTLLVATFLSKESLRMGPCSFALISKDPELQ 235
QY 393 APIRQILSRPERLLFLIDGYDE---PGWVLOEPPSSSLCHWQPOPADALLGLLGTI 448
Db 236 DDIPSILAQQRILFLVVDGDELKVPFGALIQ---DICGDWKKKVPVPLLSLKLKKM 291
QY 449 LPPASPLITARTALONLISLQARAVYVLGSSSSRRKRYFRYFDEQAIRARLVK 508
Db 292 LPPAALLVTRPRALRDLQLOAOPIYVREGFLSEBRRAVYFRHFDEQDQARAFELMR 351
QY 509 SNEELMALCLVPMWVSLACTCLMQMKRKKELTTSKTTTLTCLHYAQAOL-QAOPGPQ 567
Db 352 SNAALPOLGAPAVCVICTTKLQNEKGDPVPTCLTRGLRFLRCSFPGQAOLRGA 411
QY 568 LRDLCSLAAGIWOVKTLFSPDDLKRGDLGAIISFLKNGIIOENRIPUS-YSTHLCF 626
Db 412 LRTLSLLAAGLQWAKQSVFREDLERLGVSDELRLFLDGDILRQDRVSKGVSFIHLSF 471
QY 627 QEFPAAMSYLDEBK--RQHSNCIIDEKTEANGIHGLFGASTTR---FLGLISDE 681
Db 472 QQFLTLALFVLEKEEGEDRGHAMIDGVOKLIS--GEERLKNPDLIQVGHFLGLANEK 529
QY 682 GEEMENIFHCRLSOGNRNLQWVPSL-QLLLO-----PSLS---ESLHCLVETRN 727
Db 530 RAKELERTFCGRSS-----PDIKQELLQCKKALHANKPUSVTDLKEVLGCLYSQ 580
QY 728 KTFLTQVMAHFEBMGCVEITDMLVCTFCIKFSRHVKLQI-----769
Db 581 EELAKVAVAPFKESIHILTNTSEVMCSFSLKHQODLQKLSQVAKGVLENYMPFLDI 640
QY 770 -----IEGRQH-----775
Db 641 EFERCTYTLTPMVARODLRSLRLMTDFCSLSSNSNLKFLLEVQSFSLDSVRLCDHYT 700
QY 776 RST-----MSPTMYVL-----786
Db 701 RSTCHLOKVEIKVTPPTAVRDCCLAFIGKKTTLTLTLAGHIEWERTMLMLCDLNRNK 760
QY 787 -FRWVPV-----TDAYVQILFSLVATRNKLKELDLSGNSLSHSAVSKLCKTLRRPCLL 839
Db 761 CNQYQIRLIGHCATPEQWAEFFVYLKANGSLKTLRLSANTLDEGAMLLTKYTRPRGFL 820
QY 840 ETLRLAGCGITADDCDLAFLRANQTLTELDLSPVLTLAGAKHLQORRQSCYQL 899
Db 821 QMLSLNCRLTASCKDLAVALVSKKLTCLAKNPIGDTGVKFLCEGSLYDPCKLQTL 880
QY 900 QLVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGRVLCCEGLRHPACKLIRGLD 959
Db 881 VLQCCSTTKLGCYLSALQDAGSLTUDLSINOIAR-GIMILCOALENPNCLKHLRLK 939
QY 960 QTTLSDEMROELRALEGEKPOLLI 983
Db 940 TYETNLEIKKLBEVKEKNPKLTI 963

```

RESULT 9

```

PYAS_HUMAN STANDARD; PRT; 892 AA.
AC P59044;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 5.
GN PYPAF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;

```

```

RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Berlin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RL J. Biol. Chem. 277:29874-29880(2002).
RN [2]
RP FUNCTION.
RX MEDLINE=22275822; PubMed=12387869;
RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA Distefano P.S., Berlin J.;
RT "Functional screening of five PYPAF family members identifies PYPAF5
RT as a novel regulator of NF-kappaB and caspase-1.";
RL FBS Lett. 530:73-78(2002).
CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B.
CC -1- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
CC much lower levels in T-cells.
CC -1- SIMILARITY: Contains 1 DAPIN domain.
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL: AF479748; AAL87105.1; -
DR InterPro: IPR007091; LRR_Rninh.
DR InterPro: IPR003590; LRR_Rninh. sub.
DR InterPro: IPR007111; NACHT_NTPase.
DR InterPro: IPR004020; PAAD_DAPIN_dom.
DR Pfam: PF02758; PAAD_DAPIN; 1.
DR SMART: SM00368; LRR_R1; 3.
DR PROSITE: PS50824; DAPIN; 1.
DR PROSITE: PS50837; NACHT; 1.
KM ATP-Binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 103
FT DOMAIN 196 513
FT DOMAIN 604 614
FT REPEAT 462 487
FT REPEAT 727 747
FT REPEAT 755 778
FT REPEAT 811 834
FT REPEAT 845 868
FT REPEAT 888 909
SQ SEQUENCE 892 AA; 98733 MW; 4AA8D1FC76DDE9D CRC64;
Query Match 12.0%; Score 901.5; DB 1; Length 892;
Best Local Similarity 27.8%; Pred. No. 4.8e-50;
Matches 295; Conservative 143; Mismatches 328; Indels 295; Gaps 33;
QY 9 LACTYFLKKEIKERQOLLANKAHSSSSGET-P-AQPEKTSQMEVASLYVAQYGEQRA 66
Db 22 LLAALRELQEQKRRR-----HKLRDVGPDSRISIPWGRLEADAVLALQAFYEPBA 77
QY 67 WDLALHTWQMGRLRSICAQAOE-----GAGHSPPSPYSPSEPHLGSPPQSTAVLMW 120
Db 78 LEVARTLRRADRDVAQLQERRQLRLGLSGTLLSVSEYK-----119
QY 121 IHELPAGCTQGSERVRQLPDTSGRMREISASLLYQALPSSPDHSPQSPNAPTST 180
Db 120 -----KTKREHVLQAHAVKENASVATKTRFTLLAPF-----155
QY 181 AVLGWSGPPQSLADREQAPQTPMLDETSGIYTYEIRERERKSEKGRPPMAAVGT 240
Db 156 -----SAAPEEALGPAEPEFG-----RAERSD-----178

```

```

QY 241 PPOAHTSLOPHHHMPESVSESICSTWPMQNEDEPNQFTOLLILQRPHPRSODPLVKRSW 300
DB 179 ---THT-----FNRLFR-----188
QY 301 PDVEENRGHLIERDLFGDLPTQEPRIYILQGAIGSTLARQVKEAMGRGOLYGR 360
DB 189 ---DEGR-----RLTVLQCPAGIGTKMAKKILYMAAGKLYQCG 228
QY 361 FQHVFFSCRELAQSK-VSLAELIGDGTAPAPIRQILSRPERLLFIIDGVE-PGW 418
DB 229 VDAFEMPQCELLERPGTRSLADILIDQCPDRGAPVQMLAQPORLFIIDGDEL-P-A 286
QY 419 LQPSSELCLHMQSPQADALISGLKTLTPESSEITRTALQNLQNLISLEQARVEV 478
DB 287 LGSEAPAPCTDPEASAGAVLGLSKALLPTALLVTRAAAPGLQRLCSPOCAEV 346
QY 479 LGSSESRKEYFARYFDEERQAIRAFVLVSKNELMALCLVPWWSMLACTLMOQMRKE 538
DB 347 RGSDDKDKKRYFKFEDERRARARAYFVENETLPAICVFPVCWIVCTVLAQQLBLGR 406
QY 539 KLTUTSKTTLTCLHLYLAQLOAPL--GPQ---LRDLSLAEGIWQKTLFSPDDL 592
DB 407 DLRTSKTTLTSTVTLFTVLSAPVADGRLQGDLENLCLAREGLVGRAGFAEKEL 466
QY 593 KHGLDAIIST-FLKM-----GILQEHPIPLSYFHLCPQEFPAWSYVLE-----E 640
DB 467 QLELRGSKVQTLFLSKKELPGLVETE---VTYQFIDQSPFELALSYLLEDGVPRTAA 523
QY 641 KGRG-----KHSNCIIDELEKTLVAYGHLFGASTPFLGLISDEGERMENTFHC 692
DB 524 GGVGTILRGDAQPHSHLVL-----TTFPLGLLSAEMNRDIERHFGC 565
QY 693 RLSQ--GRNLMQV-----PSLQULLQP-----HSLESIAHCL 722
DB 566 MVSERVQKALRWVQGGQCGPVAPEVTEGAKGLDTEPEEPEEPEEPEEPEEPEEPEE 625
QY 723 YETRNKTLFLQVMAHPEKMG---CVETDMELVCTPCTKRSRAVKKQLIEGRHRS 777
DB 626 YETQEDAFVQAQCRPELALQVRFC---RMDVAVLSYCVRCPPAQALRLISCR--- 678
QY 778 TWSPVTWVLFRRWVPTDAYWQILFVSVLKTRNL---KELDSGNSLSHSAVKSICLTLR 834
DB 679 ---LVAQERKKKSLKRLQALGSSGSGGCTTKQPLASLH 717
QY 835 ---PCLLETJRLAQCGLTAEDCKDLAFGLRANQTLTELDLSPVNLTDAGAKHLC 886
DB 718 PLFOAMTDPICHLSSLTSLSHCKLPDAVCRDLSEBALRAAPALTGLLHNLSEAGRLMS 777
QY 887 QRLRQPSCKIQ--RLQ--VSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGYRL 942
DB 778 EGLAMPQCRQVQVAVQVLPDQREG---QYLVGMRLQSPALTTLLDLSGCOLPAPWVTL 832
QY 943 CEGLRHPACKLIRLGLDQTLTSDMBROELRALBEOKPOLLI 983
DB 833 CAVLQHGCGCLQTLTSLASVLSSESLQELQAVKRAKDLVI 873

```

```

RP CONCEPTUAL TRANSLATION OF 1-343.
RA Rinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RC SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
RC TISSUE=Kidney;
RC MEDLINE=96071640; Pubmed=7489366;
RA Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
RT "Identification of a novel dual angiotensin II/vasopressin receptor on
RT the basis of molecular recognition theory.";
RL Nat. Med. 1:1074-1081(1995).
RN [3]
RP VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
RX MEDLINE=21980185; Pubmed=11984003;
RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT "The dual AngII/AVP receptor gene N198/C163R variant exhibits
RT sodium-induced dysfunction and cosegregates with salt-sensitive
RT hypertension in the Dahl salt-sensitive hypertensive rat model.";
RL Mol. Med. 8:24-32(2002).
RN [4]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAP5.
RA Albrecht M., Dominguez F.S., Schreiber S., Längauer T.;
RT "Identification of mammalian orthologs associates PYPAP5 with distinct
RT functional roles.";
RL FEBS Lett. 538:173-177(2003).
CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B (By similarity). Angiotensin II and
CC vasopressin binding protein. May stimulate cAMP accumulation.
CC -1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
CC levels in all tissues tested.
CC -1- DISEASE: Defects in PYPAP5 may be a cause of salt-sensitive
CC hypertension.
CC -1- SIMILARITY: Contains 1 DAPIN domain.
CC -1- SIMILARITY: Contains 1 NACHT repeat.
CC -1- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC -1- CAUTION: The N-terminus was extended using ESTs and genomic
CC sequences, in analogy to ortholog sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M85183; AAA03623.1; ALT_INIT.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR007091; LRR_RNinh.
DR Pfam: PF00560; LRR. 2.
DR Pfam: PF02758; PAAD DAPIN; 1.
DR PROSITE: PSS0824; DAPIN; 1.
DR PROSITE: PSS0837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 103 DAPIN.
FT DOMAIN 168 484 NACHT.
FT DOMAIN 564 604 ASP/GLU-RICH.
FT DOMAIN 655 662 POLY-LYS.
FT REPEAT 433 458 LRR 1.
FT REPEAT 610 633 LRR 2.
FT REPEAT 722 745 LRR 3.
FT REPEAT 812 836 LRR 4.
FT NP BIND 174 181 ATP (POTENTIAL).
FT VARIANT 492 492 N -> R (IN SALT-SENSITIVE HYPERTENSION).
FT VARIANT 536 536 C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT MUTAGEN 766 766 E->K: ABOLISHES ANGIOTENSIN II BINDING.
SQ SEQUENCE 854 AA; 95292 MW; D7BD922D77B3734 CRC64;

```

Query Match 11.9%; Score 895; DB 1; Length 854;
 Best Local Similarity 33.2%; Pred. No. 1,2e-49;
 Matches 261; Conservative 118; Mismatches 298; Indels 110; Gaps 22;

QY 275 NOKFTQLLLQRPSPSODPLVKSMPDYEEVNRGHLIEIRDLFGPGLDTQERR--IVLL 332
 DB 114 NKRFTLLLPAGSGADELDLGTSGEPEPRARRSDTHFNRLF--KGNDEGRPLTVVL 172
 QY 333 QGAAGIGKSTLTAQVKEAMRGOLYGRFQHVHVFYFSCRELAQSK--VYSLAEILGXKDTAT 391
 DB 173 QGAGAGIKTMAAKKILYDAGAGLTHSQVPAFPMFGELLEPRGTRSLADLLLEQCPDR 232
 QY 392 PABIRQLSRPERLLFLDGVDE--PGWVLOEPPSEELCLHMSQOPADALGSLIGKTIIP 450
 DB 233 TAVRRLILAQPHRLFLILDGADLP--TLAPATPCRDPEFATSGRLVSLGSLQELLP 290
 QY 451 EASFILTAFTALQNLIPSLERQARWVVLGFSSSKREYRYRTFDRQALRAPFLVXSN 510
 DB 291 SALLVTSRATLGRLOGRLCSPQCAEVRFPSSDKKRYPFKFERKERYRFRKXEN 350
 QY 511 KETMALCLVWVWMLACTCLMOMKREKLTLSKTTTCLHYLQALQ--AQPLP--- 566
 DB 351 EITVYALCFPVPCVWICTVLLQOMELGRDLSRTSKTTTSYLLFTSMLSAGTNGBRVQ 410
 QY 567 -QLRDLCSLAEGIMOKTFLSPDDLRKGLDGAIST--FLKN-----GILQEHPIPLYS 620
 DB 411 GELRMLCRLARBGILKHQAQFSEKDLERLKLQSQVQVWFLSKKLPGVLE---TVVTYQ 467
 QY 621 FTHLCQEPFRAMSYLLEDEKGRKSNCTIIDKTLFANGIGLFG--ASTRFLILGL 678
 DB 468 FIDQSFQEFPLAALSYLDLAE--GAPCNSAGSVQMLNSDA--GLRGHLALTTRFLFGLL 522
 QY 679 SDEGEEMENIFHC-----RLSGRNLMOVPSLQILLQF----- 713
 DB 523 STERIDIGNHFCVCVPRVKKQ--DTLRWV---QGSQPVAVAVGAKKDELKDEAEER 577
 QY 714 -----HSLSLHCLYETRNKTFILYQVNAHEEMQCV-----TDMELLVCTFCT 758
 DB 578 EEEHEEEELNFGLELTYCYETQEDDFVQALSLPE--MVLEERVLFMDLEVLSTCY 635
 QY 759 KESRHKYKQLT-----EGRQHSWSPFWVLFWRVPTVDATWQILFSLKTRNL 810
 DB 636 QCCPDQALRLVSCGLVAKEKKKKKSF-----MNRLL 668
 QY 811 KELDLGNSLSHSAVSKLCTLRPRCLLETLTLAGCGLTAEDCKDLAFGRANQTLTEL 870
 DB 669 KGSQSTGKPPALPLPCAMLTQCCGSLTSLSHCKLPDANCRLSEALKAPSLREL 728
 QY 871 DLSFNVLTDAGAHLCQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDIQ 930
 DB 729 GILQNLRLTERGHLRLSQGLAMPCKVQTLRIQMPGL--QEVYIHLVLYLQSSPVLTLTLDLS 787
 QY 931 QNNLDVGVRLLEGRHAPCKLIRGLDQTLSDERQELRLLEQPKOLLIFSRKRS 990
 DB 788 GCQLPETVVERPLCSALKHPCKGLKTLSTLSVELTENPLRELQVKTLLKPDALIIHSLKLT 847
 QY 991 VMTPTG 997
 DB 848 HROPPLKG 854

RP SEQUENCE FROM N.A.
 RC STRAIN=NIN Swiss; TISSUE=Ovary;
 RX MEDLINE=99360614; PubMed=104333232;
 RA Tong Z.-B., Nelson L.M.;
 RT "A mouse gene encoding an oocyte antigen associated with autoimmune
 RT premature ovarian failure.";
 RL Endocrinology 140:3720-3726(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS ASP-7; LEU-42; LYS-87; PRO-256;
 RP GLN-977 AND 1003-ASN-ASN-1004.
 RC STRAIN=129/Sv;
 RX MEDLINE=20222768; PubMed=10754103;
 RA Tong Z.-B., Nelson L.M., Dean J.;
 RT "Water encodes a maternal protein in mice with a leucine-rich repeat
 RT domain homologous to porcine ribonuclease inhibitor.";
 RL Mamm. Genome 11:281-287(2000).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=20517328; PubMed=11062459;
 RA Tong Z.-B., Gold L., Pfeiffer K.B., Dordard H., Lee E., Bondy C.A.,
 RA Dean J., Nelson L.M.;
 RT "Water, a maternal effect gene required for early embryonic
 RT development in mice.";
 RL Nat. Genet. 26:267-268(2000).
 CC -1- FUNCTION: Necessary for embryonic development beyond the 2-cell
 CC stage.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Oocyte-specific.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF074018; AAD51762.1; -;
 DR EMBL; AF143573; AAF64393.1; -;
 DR EMBL; AF143559; AAF64393.1; JOINED.
 DR EMBL; AF143560; AAF64393.1; JOINED.
 DR EMBL; AF143561; AAF64393.1; JOINED.
 DR EMBL; AF143562; AAF64393.1; JOINED.
 DR EMBL; AF143563; AAF64393.1; JOINED.
 DR EMBL; AF143564; AAF64393.1; JOINED.
 DR EMBL; AF143565; AAF64393.1; JOINED.
 DR EMBL; AF143566; AAF64393.1; JOINED.
 DR EMBL; AF143567; AAF64393.1; JOINED.
 DR EMBL; AF143568; AAF64393.1; JOINED.
 DR EMBL; AF143569; AAF64393.1; JOINED.
 DR EMBL; AF143570; AAF64393.1; JOINED.
 DR EMBL; AF143571; AAF64393.1; JOINED.
 DR EMBL; AF143572; AAF64393.1; JOINED.
 DR PIR; A59000; A59000.
 DR HSSP; P10775; 2BNH.
 DR MGD; MGI:1345193; Water.
 DR GO; GO:0005829; Cytoplasm; IDA.
 DR GO; GO:000887; P:organelle; IMP.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR Pfam; PF00560; LRR; 3.
 DR PROSITE; PS00837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat; Polymorphism.
 FT DOMAIN 1 128 5 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 1 23 1 (INCOMPLETE).
 FT REPEAT 24 49 2.
 FT REPEAT 50 75 3.
 FT REPEAT 76 102 4.
 FT REPEAT 103 128 5.
 FT DOMAIN 191 513 NACHT.

CC - FUNCTION: May mediate activation of CASP1 via ASC and promote
 CC activation of NF-kappa-B (By similarity).
 CC - SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC - SIMILARITY: Contains 1 DAPIN domain.
 CC - SIMILARITY: Contains 1 NACHT domain.
 CC - SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
 CC - CAUTION: The N-terminus was extended using ESTs and genomic
 CC sequences, in analogy to ortholog sequences.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL, BC013519; AAH13519.1; -
 DR EMBL, BC011139; AAH31139.1; ALT_INIT.
 DR MGI, MGI:2141990; Pyraf5.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007091; LRR_RNinh.
 DR Pfam: PF00560; LRR_1.
 DR PROSITE; PS50824; DAPIN, 1.
 DR PROSITE; PS50837; NACHT, 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.
 KM
 FT DOMAIN 11 102 DAPIN.
 FT DOMAIN 168 485 NACHT.
 FT DOMAIN 569 585 ASP/GLU-RICH.
 FT DOMAIN 654 661 POLY-LYS.
 FT REPEAT 434 459 LRR 1.
 FT REPEAT 609 632 LRR 2.
 FT REPEAT 811 834 LRR 3.
 FT NP_BIND 174 181 ATP (POTENTIAL).
 SQ SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;

Query Match 11.5%; Score 868; DB 1; Length 843;
 Best Local Similarity 31.6%; Pred. No. 6,36-48;
 Matches 249; Conservative 123; Mismatches 281; Indels 136; Gaps 22;

QY 275 NQKFTQLLLQRRHPSQDPLVRSWPDYBENRGHLEIRDLF-GGGLDTQGERIYILQ 333
 DB 114 NKRFTKLILPAGTGADELLGPIGSEBEPARRSDHTNRRLFRGDESSQPLTVILQ 173
 QY 334 GAAGIGKSTARQVKAEMGRGOLYGRFQHFVYFSCRELAKSK-VSLAEILIGDGTATP 392
 DB 174 GPAGIGKTMMAKILYMAAGKLYHSQVDFAFMPCGELLERPKSLADVLVDQCPDRA 233
 QY 393 APIRQILSRPERLLFIIDGVDEPGVLOEPSEL-CLHWSQPOPADL-----LLGSLG 445
 DB 234 WPVYRIIAQENRLLFIIDGADE--LPTLPSSSEATPC-----KDLBATSGLRVLSGLTS 285
 QY 446 KTLPEASFLITATTLQNLIPSELQARWVGVGSESSRKEFYFYTFDERAIAFR 505
 DB 286 QELLPGRLLVTRHAATGRLOGLRCSPQCAELRGFSFDKCKKFFKFFEDERAAEAYR 345
 QY 506 LYSNKEMLALCLVPWVSWLACTCLMQOMRKKEKLTLSKTTTLTCLHYLAQALQ-AQPL 564
 DB 346 FVKNENLPLALCFPFCWCIVCTVLOOQLLEGRPLSTSKTTISVYLFTLSMLKSGTN 405
 QY 565 GP-----QLRDLCSIAAGIWM-QKTLFSPDLRRHGLDGAIIISFLKMGILQENRIP 616
 DB 406 GPRVQGLRLLCLARREGIILDHNAQFSEEDLETKLRGSGVQTI-----FLANKKEIRGVL 461
 QY 617 ---LSYSEFHLCPQEFPAFAMSVYLEDKSGKHSNCIIDEKTEAVYGINGLPG----- 667
 DB 462 KTEVTVFIQDSFQEFPLAALSYLSEAKETPP-----TPAGGVQKLSMDAELR 509
 QY 668 ---ASTRFLGLSDSGEREMENIFQRLSQ--GRNLMQWVPSLQULLQPH----- 714
 DB 510 GHLLATTRFLFGLNLTGRLDIGNHFGCVVVDHVKQDRLRVVQG-----QSHPRGPPVGA 564

QY 715 -----SLESJLCIVETRNKTFELQVMAHFEEEMCVB----- 746
 DB 565 KKTAALEDIEDABEEBEEEDLNFGLELLYCLVETQEDPVRQALSLPE--IVLERVRL 622
 QY 747 TDBELLVCFICIFSRSHVKKLOI-----BGRQHRSTWSTPMVLPFWVVTDAVWQ 798
 DB 623 TRNDLELVNLCVQCCPDGQALRLVSCGLVAANEKKKKKKK----- 662
 QY 799 ILRSVLKVTNRLNELDLSGNSLSHSAVSKITLRAPRCILLETLRAGCGLTAEDCKDLA 858
 DB 663 -----LVRLKSGSGSTKQPPVSLRLPLCEMTTPKCLSTVILSHCRLPDAVCRDLS 715
 QY 859 EGRANQTLTELDLSFNVLTDAGAKLQRLQPSCKLQRLQVSCGLTSDCCQDLASVL 918
 DB 716 EALKVAPALRELDLQSRLLNTGLRLCEGLAMPKQCVKTLRMQLPDL-QEVNLYVIVL 774
 QY 919 SASPSLKELDLQNNLDVGVRLCGSLHHPACKLRLGLDQTLTSDERQELRALEQK 978
 DB 775 QQSPVLTTLDDLSCQPLPVATVEPLCALHHPKCSLTSLTSVLSBNSLRDLQAVKTSK 834
 QY 979 POL-LIFSR 986
 DB 835 PDLSTYSK 843

RESULT 13
 ID PYA6_HUMAN STANDARD; PRT; 1033 AA.
 AC P59045; O8NBFS.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 6.
 GN PYPAF6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Bilekin M., Distefano P.S., Bertin J.,
 RA "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing";
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang U., Heien F.,
 RA Diatchenko L., Marushia K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,
 RA Bampstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
 RA Vallat D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smilins D.E.,
 RA Schmeich A., Schein U.E., Jones S.J.W., Natta W.A.,
 RT "generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RN SEQUENCE OF 365-1033 FROM N.A. (ISOFORM 2).

RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21950691; PubMed=11821383;
 RA Razmara M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
 RA Distefano P.S., Bertin J., Alnemri E.S.;
 RT "CARD-8 protein, a new CARD family member that regulates caspase-1
 RT activation and apoptosis.";
 RL J. Biol. Chem. 277:13952-13958(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Zhang H.;
 RT "A novel apoptotic protein, NDDP1, containing CARD and BH3 domains.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Guet C., Vito P.;
 RT "DACCAR, a novel CARD-containing protein.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21570185; PubMed=1151959;
 RA Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
 RA MacFarlane M., Martin S.J.;
 RT "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
 RT of multiple NF-kappa B activation pathways.";
 RL J. Biol. Chem. 276:44069-44077(2001).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC TISSUE=Kidney;
 RA Guo J.H., Yu L.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE=21402909; PubMed=11408476;
 RA Padhnan N., Marusawa H., Krajewska M., Mateuzawa S.-I., Kim H.,
 RA Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Plo F.,
 RA Godzik A., Reed J.C.;
 RT "TUCAN, an antiapoptotic caspase-associated recruitment domain family
 RT protein overexpressed in cancer.";
 RL J. Biol. Chem. 276:32220-32229(2001).
 RN [8]
 RP CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
 RX MEDLINE=22062958; PubMed=12067710;
 RA Stillo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
 RT "TUCAN/CARDINAL and DRA1 participate in a common pathway for
 RT modulation of NF-kappa B activation.";
 RL FEBS Lett. 521:165-169(2002).
 CC -1- FUNCTION: Inhibits NF-kappa B activation. May participate in a
 CC regulatory mechanism that coordinates cellular responses
 CC controlled by NF-kappa B transcription factor. Involved in the
 CC negative regulation of caspase-1.
 CC -1- SUBUNIT: May form homodimers. Interacts with NEMO and DRA1. Binds
 CC to caspase-1, pseudo-ICE and ICEBERG.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q9Y2G2-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q9Y2G2-2; Sequence=VSP_000782, VSP_000783;
 CC -1- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
 CC placenta. Lower expression in heart, kidney and liver. Also
 CC expressed in spleen, lymph node and bone marrow.
 CC -1- SIMILARITY: Contains 1 CARD domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL, AB023172; BAA76799.1; -.

DR EMBL, AF322184; AAG50014.1; -;
 DR EMBL, AF331519; AAK01126.1; -;
 DR EMBL, AY026322; AAK08982.1; -;
 DR EMBL, AF405558; AAL02427.1; -;
 DR EMBL, AF511652; AAM46959.1; -;
 DR InterPro: IPR001315; CARD.
 DR Pfam: PF00619; CARD; 1.
 DR SMART: SM00114; CARD; 1.
 DR PROSITE: PS0209; CARD; 1.
 KW Apoptosis; Nuclear protein; Alternative splicing.
 FT DOMAIN 340 430
 FT VARSPPLIC 282 286
 FT CARD.
 FT ELKS -> MISL (in isoform short).
 FT VARSPPLIC 287 431
 FT /FTID=VSP_000782.
 FT MISSING (in isoform short).
 FT MUTAGEN 366 366
 FT L->R: INHIBITS HOMODIMER FORMATION.
 FT E->G (IN REF. 5).
 FT V->M (IN REF. 5).
 FT L->P (IN REF. 5).
 FT CONFLICT 422 422
 FT SEQUENCE 431 AA; 48932 MW; C854D1308073286 CRC64;
 Query Match 8.9%; Score 668; DB 1; Length 431;
 Best local similarity 39.5%; Pred. No. 1.7e-35;
 Matches 151; Conservative 53; Mismatches 122; Indels 56; Gaps 7;
 QY 1079 FNGPTGPVATEVVDKRNLYRHPVAGSYKRNPNGLCFVMBAYVVEIFCVWDQFLG- 1137
 DB 56 FLGPBEGNVDELIDKSTNRYSVWFPAGWYLSATGLGLVDEVTITAFGSWSCHLAL 115
 QY 1138 EINPQSMNVNAPPLDILKAP-GAVEAVHLPHFVALQGHVDTLSQMAHFKREGMLTK- 1196
 DB 116 DLQHEHQWLVGGFLPVTAPEBAVAEHLPHFISLQ-GEVDVSWFLVHFKNEGVLTH 174
 QY 1197 PARVELHIVLENPSSPGLVLLKMINALRFIPVTVSVLYLHYRVBEEVTHLYLIPSD 1256
 DB 175 PARVEPFVAVLSPSSLSMGLILRLASGRRLIPITSNTLIYHPHPEDIKFHLVYVPSD 234
 QY 1257 CSIR-----KELELCRSGEQ 1274
 DB 235 ALLTKAIDDEDRFHVRLQTSPPMEPLFGSSYIVSNANLKWPKELKLSRSGEIQ 294
 QY 1275 LPSEPFVGHLSGIRLOVNDKDETLWEALVKPGDLPATLLIPARLAVSPPLDAPQL 1334
 DB 295 HSKPYAGMKPQLQLEITEKKGITLVDTKYKVDQLVAASAP-----PFGAA- 346
 QY 1335 LHFVDQYRQLIARYTSVEVLDKLN-GQVLSQOYERYVLAENTRPSQMKLFSLSQSWD 1393
 DB 347 --FVKENHRLQIARWGDLDKGLVDDLDQDNEVLITENEKELVGEKTRQS KNEALLSWETKG 404
 QY 1394 RKCKDGLYQALXETHPHLIMEL 1415
 DB 405 DLALDVLFRSISBRDPYLVSYL 426
 RESULT 15
 CARD_HUMAN STANDARD; PRT; 1040 AA.
 AC Q9HRC29; Q96RH5; Q96RH6; Q96RH8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-FEB-2003 (Rel. 42, Last annotation update)
 DE Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory
 DE bowel disease protein 1).
 GN CARD15 OR NOD2 OR IBD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND
 RP VARIANT ARG-908.
 RC TISSUE=Breast;
 RX PubMed=11087742;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 79.1934 Seconds
(without alignments)
4656.416 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGANGRACLYEFLKKE.....HLIMEIMKSKKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL.23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2640	35.0	509	4 Q9H5Z8	Q9H5Z8 homo sapien
2	1754	23.3	409	4 Q9H5Z7	Q9H5Z7 homo sapien
3	1371.5	18.2	882	4 Q8NEU4	Q8NEU4 homo sapien
4	1254.5	16.7	1033	11 Q8R4B8	Q8R4B8 mus musculi
5	899.5	11.5	982	11 Q8BU40	Q8BU40 mus musculi
6	868	11.5	863	11 Q8C6J9	Q8C6J9 mus musculi
7	843	10.8	846	4 Q8IXT0	Q8IXT0 homo sapien
8	815.5	10.8	713	6 Q95LZ7	Q95LZ7 macaca fasc
9	689.5	9.2	748	11 Q99MM0	Q99MM0 mus musculi
10	679.5	9.0	825	11 Q8C6M5	Q8C6M5 mus musculi
11	676	9.0	657	11 Q9EPG7	Q9EPG7 mus musculi
12	573	7.6	673	11 Q8CCN1	Q8CCN1 mus musculi
13	393.5	5.2	953	4 Q8IMF5	Q8IMF5 homo sapien
14	392	5.2	1020	11 Q8K3Z0	Q8K3Z0 mus musculi
15	383	5.1	287	4 Q9BYZ6	Q9BYZ6 homo sapien
16	372.5	4.9	778	4 Q8NFA8	Q8NFA8 homo sapien

17	363.5	4.8	456	11 Q91V17	Q91V17 mus musculi
18	361.5	4.8	456	11 Q924P4	Q924P4 mus musculi
19	358	4.8	447	4 Q96FD7	Q96FD7 homo sapien
20	358	4.8	461	4 Q9BQ80	Q9BQ80 homo sapien
21	357	4.7	461	6 Q8HZP9	Q8HZP9 pan troglod
22	351.5	4.7	390	11 Q8BWZ1	Q8BWZ1 mus musculi
23	351	4.7	953	11 Q8BHB0	Q8BHB0 mus musculi
24	348	4.6	461	4 Q81ZK8	Q81ZK8 homo sapien
25	339.5	4.5	516	4 Q8NBP5	Q8NBP5 homo sapien
26	334	4.4	1052	7 Q9GJD8	Q9GJD8 ratu mus norv
27	334	4.4	1073	7 Q9GJD9	Q9GJD9 ratu mus norv
28	334	4.4	1153	7 Q9GJBO	Q9GJBO ratu mus norv
29	325	4.3	269	11 Q9D458	Q9D458 mus musculi
30	316.5	4.2	1155	7 Q9TPE1	Q9TPE1 mus musculi
31	316.5	4.2	519	11 Q8C249	Q8C249 mus musculi
32	304	4.0	660	4 Q8NPF6	Q8NPF6 homo sapien
33	291.5	3.9	733	4 Q8TEE2	Q8TEE2 homo sapien
34	289	3.8	977	7 Q8HW99	Q8HW99 mus musculi
35	280.5	3.7	884	7 Q26755	Q26755 homo sapien
36	278.5	3.7	932	4 Q96KL4	Q96KL4 homo sapien
37	277	3.7	1056	4 Q8NPF42	Q8NPF42 homo sapien
38	276	3.7	692	4 Q96D51	Q96D51 homo sapien
39	272.5	3.6	706	11 Q8BUT6	Q8BUT6 mus musculi
40	258	3.4	1087	4 Q9HXY0	Q9HXY0 homo sapien
41	241	3.2	195	6 Q8HXK9	Q8HXK9 bos taurus
42	232.5	3.1	1402	11 Q8CH65	Q8CH65 mus musculi
43	232.5	3.1	1403	11 Q8CH70	Q8CH70 mus musculi
44	226.5	3.0	1402	11 Q8CCT4	Q8CCT4 mus musculi
45	223.5	3.0	1403	11 Q8BGB6	Q8BGB6 mus musculi

ALIGNMENTS

RESULT 1

Q9H5Z8	PRELIMINARY;	PRT;	509 AA.
AC Q9H5Z8;			
DT 01-MAR-2001 (TREMBLrel. 16, Created)			
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DE Hypothetical protein FLJ22740.			
OS Homo sapiens (human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RP [1]			
RA KAWAKAMI T., NOGUCHI S., ITOH T., SHIGETA K., SENBA T., MATSUMURA K.,			
RA YAMADA K., FUJII Y., MIZUNO T., MORINAGA M., TANIGAMI A., FUJIWARA T., ONO T.,			
RA OBAHASHI M., NISHI T., SHIBAHARA T., TANAKA T., NAKAMURA Y.,			
RA ISOGAI T., SUGANO S.;			
RT "NEO human cDNA sequencing project."			
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.			
DR EMBL; AK026393; BAB15469.1; -			
DR HSSP; P10775; 2BNH.			
DR InterPro; IPR001611; LRR.			
DR InterPro; IPR007091; LRR_RNinh.			
DR Pfam; PF00560; LRR_1.			
DR PROSITE; PS50503; LRR_R1; 2.			
DR Hypothetical protein.			
SQ SEQUENCE 509 AA; 56908 MW; EB7535AF69817F5B CRC64;			

Query Match 35.0%; Score 2640; DB 4; Length 509;

Best Local Similarity 100.0%; Pred. No. 2.2e-190; Indels 0; Gaps 0;

Matches 506; Conservative 0; Mismatches 0;

QY	735 MAHFEEMGCVETDMLVCTFCIKSRHYKQLQILBGRQHSSTWSPVWVLEFWVPVTD	794
DB	1 MAHFEEMGCVETDMLVCTFCIKSRHYKQLQILBGRQHSSTWSPVWVLEFWVPVTD	60
QY	795 ATWQILFVLTAKYTRNLKEIDLSGNSLSHSAVSLCKTLRRPCLLETLLAGGLTAEDC	854

```

Db      61  ATWQILFSAVKTRNKLELDLSDNSLSHSAVSLCKTLRPRCLLETLRLAGCGLTAEDC 120
QY      855  KDLAFGLRANQTLTELDLSENVLTLDAGAKHLCORLRQPSCKLQRLQVSCGLTSDCCDL 914
Db      121  KDLAFGLRANQTLTELDLSENVLTLDAGAKHLCORLRQPSCKLQRLQVSCGLTSDCCDL 180
QY      915  ASVLSASPSLKEIDLQNNLDVGVRLCEGLRHAPCKLIRGLDQTLTSDENRQELRAL 974
Db      181  ASVLSASPSLKEIDLQNNLDVGVRLCEGLRHAPCKLIRGLDQTLTSDENRQELRAL 240
QY      975  EOEKPOLLFRRKPSVMTPTBGLDTGEMSNSTSLKROLGSRASASHVAQNLKLLV 1034
Db      241  EOEKPOLLFRRKPSVMTPTBGLDTGEMSNSTSLKROLGSRASASHVAQNLKLLV 300
QY      1035  SKIFPFAETAEBSSEPVVVELLCVPSASQGLHTKPLCTDDFPGPTGPVATEVVDKE 1094
Db      301  SKIFPFAETAEBSSEPVVVELLCVPSASQGLHTKPLCTDDFPGPTGPVATEVVDKE 360
QY      1095  KNLVYHFPVAGSYRNPNTGLCFVMEBAVTEIEFCVWDQFGEINPQSHMVAAGPLDI 1154
Db      361  KNLVYHFPVAGSYRNPNTGLCFVMEBAVTEIEFCVWDQFGEINPQSHMVAAGPLDI 420
QY      1155  KAEFGAVEAHLPHFVALQGGHDTSLFQMAHKEBGMLEKPARVELHIVLENSFSP 1214
Db      421  KAEFGAVEAHLPHFVALQGGHDTSLFQMAHKEBGMLEKPARVELHIVLENSFSP 480
QY      1215  LGVLLKMINALRFIPVTSVLLYHR 1240
Db      481  LGVLLKMINALRFIPVTSVLLYHR 506

```

RESULT 2

```

Q9HSZ7  PRELIMINARY; PRT; 409 AA.
AC      Q9HSZ7;
DT      01-MAR-2001 (T-EMBLrel. 16, Created)
DT      01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT      01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE      Hypothetical protein FLJ22745.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA      Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujimura T., Oho T.,
RA      Yamada K., Fujii Y., Ozaki K., Hideo M., Ohmori Y., Ota T., Suzuki Y.,
RA      Odayashi M., Nishii T., Shibahara T., Tanaka T., Nakamura Y.,
RA      Isogai T., Sugano S.;
RT      "NEO human cDNA sequencing project.";
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK026398; BAB15470.1; -
KM      Hypothetical protein.
SQ      SEQUENCE 409 AA; 46068 MW; 07CC5FACF3EB736 CRC64;

```

```

Query Match      23.3%; Score 1754; DB 4; Length 409;
Best Local Similarity 87.7%; Pred. No. 1,1e-123;
Matches 343; Conservative 0; Mismatches 0; Indels 48; Gaps 2;

```

```

QY      967  MROELPALOEKFOILIFSRKPSVMTPTBGLDTGEMSNSTSLKROLGSRASASHVAQ 1026
Db      1  MROELPALOEKFOILIFSRKPSVMTPTBGLDTGEMSNSTSLKROLGSRASASHVAQ 60
QY      1027  ANKLKLDVSKIFPIAETAE-----EBSSEPVVVELLCVPSASQGLHTKPLCTDDFPG 1082
Db      61  ANKLKLDVSKIFPIAETAEBSSEPVVVELLCVPSASQGLHTKPLCTDDFPG 120
QY      1083  TGEVATEVVDKEKNLYRHFVPAVSGYRNPNTGLCFVMEBAVTEIEFCVWDQFGEINQ 1142
Db      121  TGEVATEVVDKEKNLYRHFVPAVSGYRNPNTGLCFVMEBAVTEIEFCVWDQFGEINQ 180

```

```

QY      1143  HSMVAAGPLDIKAEFGAVEAHLPHFVALQGGHDTSLFQMAHKEBGMLEKPARVEL 1202
Db      181  HSMVAAGPLDIKAEFGAVEAHLPHFVALQGGHDTSLFQMAHKEBGMLEKPARVEL 240
QY      1203  HHIIVLENSFSPGLVLMKMINALRFIPVTSVLLYHVFHEVVEFHYLIPSDSIR-- 1260
Db      241  HHIIVLENSFSPGLVLMKMINALRFIPVTSVLLYHVFHEVVEFHYLIPSDSIR 300
QY      1261  -----KELELCYRSPGEDQLFSE 1278
Db      301  IDDLKMKQFVAIRHKPPLTPLYMGCRYTVSGSGSGMELIPKELELCYRSPGEDQLFSE 360
QY      1279  FVYGHGSGIRLQYOKXOKDETLWEALYKPG 1309
Db      361  FVYGHGSGIRLQYOKXOKDETLWEALYKPG 391

```

RESULT 3

```

Q8NEU4  PRELIMINARY; PRT; 892 AA.
AC      Q8NEU4;
DT      01-OCT-2002 (T-EMBLrel. 22, Created)
DT      01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE      Monarch-1 splice form IV.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Williams K.L., Linhof M.W., Ting J.P.Y.;
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY116207; AAM75145.1; -
DR      InterPro: IPR007091; IRR RNinh.
DR      InterPro: IPR007111; NACHT NTPase.
DR      InterPro: IPR004020; PAAD_DAPIN_dom.
DR      Pfam; PF02758; PAAD_DAPIN; 1.
DR      PROSITE; PS00824; DAPIN; 1.
DR      PROSITE; PS00503; IRR RI; 1.
DR      PROSITE; PS00337; NACHT; 1.
SQ      SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

```

```

Query Match      18.2%; Score 1371.5; DB 4; Length 892;
Best Local Similarity 35.2%; Pred. No. 2,6e-94;
Matches 356; Conservative 135; Mismatches 349; Indels 171; Gaps 22;

```

```

QY      8  RLACTLEPLKKELEKEFOLLANKHSSSGSETP-AQPEKTSMEVASIYVAQGEQRA 66
Db      12  RLSTLELELEAVALKFKFLYLGTA--TELGEKTIWGSMEKAGPLEMQLLITHFGEPEA 69
QY      67  WDLALHTWEQWLRLSLCAQAGEAGHSFPSPSEPHLIGSPSQ-PTSTAVLMPWIHEL 125
Db      70  WDLALSTFERIRIKOLMGRGREDLVROTPGSPSS--LGNGSTGLLEVSUTP--RKDP 125
QY      126  AGCTGSGSERVRLQPLDTSGRMRREISASLLYQALPSPDHSPQESPNAPTSTAVALGS 185
Db      126  QETRYDYARRKFRKLMDENARLGECVNLSHRYTRLLVYEHSNPQVQ----- 173
QY      126  WSPSPQPSLAPRQAPRGQWLDTSGLTYEIERERERESEKRRPMAAVVGTRPQAH 245
Db      174  ----- 173
QY      246  TSLQPHHHPWEPSVRESICSTWPMXNEDFNQKFTOLLLOPHRPSODPLVRSWPDVVE 305
Db      174  -----QQLLDTGRGHATV----- 187
QY      306  EKRGLH---IEKRDLPGEOLDTQF-PRIVILQAGAGIGKSTLARYKQKAWGGLYGRDF 361
Db      188  ---GQASPIKLETLFEEDERPERPRVVMQGAIGKSMIAHAKVMDWADGKLFQGRF 244
QY      362  QHVFESCEELASQKV-VSLAELIGKDGATATAPARQLISREERLLFILDGDEGVWLQ 420
Db      420  QHVFESCEELASQKV-VSLAELIGKDGATATAPARQLISREERLLFILDGDEGVWLQ 420

```

Db 245 DYLFYINCRMNQSGATECSMODLIFSCEPESAPLOELIRVERBLFIIDGDELDKPSFH 304
 QY 421 EPSESELHNSQOPADALLGSLGKTLTPASFLITRTALONLIPSLQAWVEYLQ 480
 Db 305 DPQGPWLCWEKPEKTELLNSLRKCLPBLSLITTRPALREKRLLEHPHVEILG 364
 QY 481 FSESSRKEFYFRYPTDERQAIIRAFRLVKSNEKLMALCVPMVSMIACGLCMQKREK 540
 Db 365 FSEERKEFEYFKYTHNAEQAGVRYVRDNEPLTMCVPLVCVAVCTCLOQOEGGGL 424
 QY 541 TLTSKTTTLCIHYLAQALQAPLPOL-----RDCLSLAEGSIWOKTLPSPDRLKH 594
 Db 425 RQTSRTTAAVYMLYLSIMQKPGAPRLQPPNQGCSLAADGLMNOKTLFEEQDLRKH 484
 QY 595 GLDPAITSTLKMILQEH-PIPLSYSTIHCPOEPFPAANSYVL-EDKRGKRSNCIID 652
 Db 485 GLDDEDSAFPLNMNIPOKINCERYSTIHSFQEFPAAMYIIDEBSGGAGPQ----D 540
 QY 653 LEKTLAAYGI-HGFLGASTTRFLGLLSDGEREMENIFHCRLSQ--GRNLMQWVSLQ- 708
 Db 541 VTRLLTYARSEERFLALTSRFLGLNEETRSHLESKLVKVPPIKMDLQWIOQKAO 600
 QY 709 ---LLLOPHSLSLHCLYETENKFTLTQVNAHPEEMGMC-VETIMELVCTFCIKFSRHV 764
 Db 601 SDGSTLQGGSLFEPSCLYEIEEEFIQALSHFQIVVSNINASIMHNVSFCLKGRSA 660
 QY 765 KKLQOLI-----EGQHSSTWPTVYVLR--WVPYDAIWOILFSLVKTNLKYLD 814
 Db 661 QVHLVYATYSADDEBARCSAGAHHTLVORPERTVLLDAYSHEHLAALCTNNILSL 720
 QY 815 LSGNSLSHSAVKSCLKTLRPRCLLETLRLAGCLTAEDCKDLAFGRANOTLLELDSF 874
 Db 721 LYRNALSRGVKLLCOGLRHPNCKLONLRKCRISACEDLSAAILANLNLRMDISG 780
 QY 875 NVLTDAKAKHLCQRLRQPSCKLQRLQVLSGCLTSDCCODLASVLSASPSLKELDLQONNL 934
 Db 781 NGVGFPGMWLLCEGLRHPQCLQWIOLRKQLESAGQEMASVLTGPHVLEDDLGNAL 840
 QY 935 DDVAVRLCEGLRHPACKLRL---GLDQTTLSDEKQELRALQEKPOL 961
 Db 841 EDGLRLCLCOGLRHPVCRRLRTLMWLFGMD--LNKMTSHSLAALRVKPYL 888
 RESULT 4
 QY 08R4B8 PRELIMINARY; PRT; 1033 AA.
 AC 08R4B8;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Mast cell maturation inducible protein 1.
 GN MMIGL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/cJ;
 RA Kikuchi-Yanoshita R., Koga K., Takeomi Y., Sugiki T., Saito T.,
 RA Ichih S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
 RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo T.,
 RT "Identification of inducible genes during in vitro maturation of mouse
 RT bone marrow-derived mast cells to connective tissue-type mast cells."
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF486632; AAL90874.1; -
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001611; LRR_1.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.

DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_R1; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS50503; LRR_R1; 3.
 DR PROSITE; PS50837; NACHT; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR SEQUENCE 1033 AA; 118274 MW; 592469066B12117 CRC64;
 Query Match 16.7%; Score 1254.5; DB 11; Length 1033;
 Best Local Similarity 32.6%; Pred. No. 2,2e-85;
 Matches 331; Conservative 164; Mismatches 334; Indels 187; Gaps 23;
 QY 8 RLACVLEFLKEEKEKEROILLANKAHSSSSGEPNQPKEKTSQWNEVASYVAQGEORAW 67
 Db 7 KLAQVLELDLDVDLKKFKMLBDFPEKGCIPVPRGQWEDADHDLDLWIDENGEKAW 66
 QY 68 DLALHTWQWGLSLCAQAOEGAGHSFPSPSEPHLGSPOPTSTAVLPIWHELPAQ 127
 Db 67 AMAVWIFPAINRRDIMEKAKK-----DQP-----EWNDT 95
 QY 128 CTQGSERRVRLQPLDTSGRWRREISASLYQALPSPDHSPSQESPNAPTSTAVLQSG 187
 Db 96 CTHSSS-MVCOE--DSLSEEM-----MGILG 118
 QY 188 SPQPSLAREQAPRGQWLDTSGLTYTEIREREREKSEKGRPPAAVGTTPQAHNTS 247
 Db 119 YLSRISCTCKKK-----DYCKRYRRHVSFRYSIKRN-----ARLG----- 155
 QY 248 LQPHHHPWESVRESLCTWPKNEDENCKFTOLLQRPHPRSQDPLVRSWPDVVEEN 307
 Db 156 -----ESVDLSRYTQQLV-KHPSKQE----- 178
 QY 308 RGHLI-----EIRD-----LFGPLDQEP-RIVLQAGAGIGSTARQVKEA 350
 Db 179 REHELLITIGTKWRDSPMSLKLLELFEPEPDGSEPHYTVFOGAAGIGITILARKMLD 238
 QY 351 WGRGQLYGDRFGVVFVYSSCELAQSKVSLAEILGKGTTPAPIRQILSRPBLFIID 410
 Db 239 WALGKLFKDFYDLFFHCRVSLRTPRSIADLIVSCWDPNPDPVCKILRPSLILMD 298
 QY 411 GVDPEGVLOEPPSESLCLHNSQOPADALLGSLGKTLPEASEFLITARTALONLIPSL 470
 Db 299 GFDELQGFADBEHIGEVCTDQKAVRGDILLSLIRKCLLPKASLLITTRVPALEKQHL 358
 QY 471 EQARWEVLCFSSSRKEFYRYFTDERQAIIRAFRLVKSNEKLMALCVPMVSMIACGL 530
 Db 359 DHPHVEILGFSEAKRKEFYFKFSNELQARBAFRLQENEVLFQWCFIDPLVCWICTGL 418
 QY 531 MOQMKREKTLTKSKTTTLCIHYLAQALQAP-----LGPQLRDCSLAABSIWOKT 584
 Db 419 KQWETGKSLAOTSCTTAVYVFFLSLLQSRGIEHILFSDYQGLCSLAADSIWOKI 478
 QY 585 LPSDDLKRGKGLQALISTFLKGILO-EHPILSYSTIHCPOEPFPAANSYVLEDE--- 640
 Db 479 LPFECDLRKGKGLQKTDVSATLRMNVFQKVEDCERFYSFSHMTQEPFAANYVILLEEAG 538
 QY 641 -----KGRKRSNCI-IDLEKTLAAYG-IGLFGASTTRFLGLLSDGEREMENIFHCRL 694
 Db 539 ETVRKGGGCGSDLLNRDVKVLLLENYGFKEKGYLIFVVRFLGLVNOQRTSYLEKSLCKI 598
 QY 695 SQ--GRNLMQW-----PSIOLLQPHSLSLHCLYETENKFTLTQVNAHPEEMGMCVETD 746
 Db 599 SQQVRLLELLEIVKAKAKLQWQPSQLFPCLYEQEBSFVQSANMDFPKLEINLSTR 658
 QY 749 MELVCTFCIKFSRHVAKQLQ-----LEGROHSTWSPF-----WVVLFRVY 790
 Db 659 MDHVSFFCIKNCHRYVTLSLGFPHNSPKKEEERGRGLDQVQCVFPTTHVACSSRLV 718
 QY 791 P--VTDAIWOILFSLVKTNLKYLDLSGNSLSHSAVKSCLKTLRPRCLLETLRLAGCG 848
 Db 719 NCCITSSFCGLFSSLSITNSLTLELDSNTLDPGRVRLCEALQHGHCNIORLMLRCGC 778
 QY 849 LTAEDCKDLAFGRANOTLLELDSFVNLTDAGAKHLQGLRQPSCKLQRLQVLSGCLTS 908

```

DB 779 LSHQCCFDLSSVSSSQKLELDLSDNALDPEGRLLCVGLKLNLCQMLVSSCLTS 838
QY 909 DCCDIALSVLSAPSLKELDLQONNLDVGVRLLCGLRHPACKLIRLGIDQTTLS 964
DB 839 ACCODIALVLSNHSILRLYIGENALDGSVGLCEMRKDPQCNLQKLGIVNSGLT 894

RESULT 5
Q8BU40 PRELIMINARY; PRT; 982 AA.
ID Q8BU40
AC Q8BU40;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to PAN2 protein.
DS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Ovary;
RX MEDLINE=2354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RL EMBL; AK087843; BAC40024.1; -.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFEA2053585 CRC64;

Query Match 11.9%; Score 899.5; DB 11; Length 982;
Best Local Similarity 26.8%; Pred. No. 1.2e-58;
Matches 269; Conservative 106; Mismatches 366; Indels 191; Gaps 26;

QY 203 GTCWPLDENSGIYVIRE-REKSEKSG--RPMVAVGTPTPQATSLQPHH-----PW 255
DB 9 GLMMYIEELNKKKEFVFKFELKQEVLTQGLKQSWTEVKKASQDLASLLKHYEKPRM 68
QY 256 EBSVRESLSTWPKNEDFNQKFTQLLLQRP-----PRSDPLVKTSWPDYVENRG 309
DB 69 DMTFR-----PFQKINRDLIERAKELIDGCKLYRAMKTKMTH--DSSRA 113
QY 310 HLIEIR-----DLFGGLDTQEPRIYVILQGAAGIKSTIAROVKAWGR 353
DB 114 FTISIQNPLKEKETEDDYDCFNLFQSKGTESKQVVFISGAGVGTMLTKRIMLAWIE 173
QY 354 GQYGGRRQGVHVFYFSGRELAQSKVSLAELIGDGTATPAPRIQILSRPRLIFILGDV 413
DB 174 SPVFLHFKFSYIFCCREVKQLTASIAELISREWPQSPAPRIEILSKPEKLLFITDSIE 233
QY 414 EPGWVQEPSSSELCLHWSQPOPADALIGSLGTLIPASFLITARTTALQNLIPSLQGA 473
DB 234 GMECDLFFKMESELCDNTEKQPNVNLISLRKMLPESLLISAPRESKEMENIYRT 293
QY 474 RWVEVLGFSSSSKKEYRYFTDERQAIRAFLVKSKEIMALCLVPWWSWLACTLMOQ 533
DB 294 HWKILIGLKERNIKMFHRLFQDRNRAHEAFSLVRENEQLFTVCQVPVLCMVATCLKEB 353
QY 534 M-KRKREKLTITSTKTTTLCIHYLAQLQAPLGR-----QLADLSIAAGIWMQKTLF 566
DB 354 IEKGRPVVSCRCCTSLYTHHNLFPONAHSPSKSDODLQGLSLAEGWMTDTPVF 413
QY 587 SPDDLKRGHGLDGAIIITFLMGIL--QBNPIPLSYFTHLCFOEPFAMSVYLEDKGRGK 645
DB 414 GSEALARRNGIMDIDPILDLIGLINRESEKSYIFLHPVGVCALIFLL--KSHVD 470
QY 646 H-SNCIIDLKTYEAY---GIGLFGASTTRFLGLSLDEGERMENTHCLSLQ--GR 698
DB 471 HPSQEVKSEIKMAFAFLKAVKQVWIFPGS---FIFGLHSEQKLEAFPGHLSQSEIKR 527
QY 699 NLMQWVPSL-----QLLLQPHSLSLHCLVETRNKTFITQVMAFPEMGWVETDMELVLC 754

```

```

DB 528 QLYQCLFETISGBEBLEQIDGKMLFYCLPEMDDDFTLVANMCMEOINVAADYSVITA 587
QY 755 TFCIKFSRRVKKL-----QLIEGRQHS-----TWSEPTVNL-----FRWVEV 792
DB 588 AACLKHCFTLKKLSTSTQVNLGSAQHSYMERLLTAMNHI CSVFISKDIQELRMKDITNL 647
QY 793 TDAYWOILFSLVKTR-----NKELDLSGNSLSHSAY 825
DB 648 SESAFSVLYNNLKRYNHTLVANNVFPVCEKYLFELLIONCNLOHMLSLTILSHSDV 707
QY 826 KSLCKTLRRPCLLETTLRAGGGLTAEDCKDLAFGRANQTLTELDSPNVLTADAKHL 885
DB 708 KLLCDVLSQACNTEBELVVAACSISDDCKVAVSLISNKLTKHNLASNTL-DKGIASL 766
QY 886 CQRLRPSCKLQRLQVSGCLTSDC-----910
DB 767 CSLCHPDCILHVLVANGSLNBEKMDVLSVLRNKKTLSHLDISSNDLKDEGLKYLGRA 826
QY 911 -----CODLASVLSAPSLKELDLQONNLDVGVRLLCGLR 948
DB 827 LTLPLDSVLSLSLRCLTITISGCDLAEVLRNNQNLVLSQVNNKLEDTGVLLCDALRH 886
QY 949 PCKLIRIGLDQTTLSDEMROELRALDEKPOLLIFSRRKPSVMTPEGLDTGEMSNSTS 1008
DB 887 PCHLIEDLGLBACEILGACCEDLASTFTQCKTLMAANLIK-----NALDYNGLVLCB 939
QY 1009 SLKQRLGSERAASHVAQNLKLLDV-SKIPRIAEIABESSP 1049
DB 940 ALKQ-----TQATVYLGQIITDPDETQAFVLAE--QEKNP 974

RESULT 6
Q8C6J9 PRELIMINARY; PRT; 863 AA.
ID Q8C6J9
AC Q8C6J9;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical RNI-like structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Ovary;
RX MEDLINE=2354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RL EMBL; AK054426; BAC35775.1; -.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 99955 MW; D3A46C782DA544CF CRC64;

Query Match 11.5%; Score 868; DB 11; Length 863;
Best Local Similarity 27.7%; Pred. No. 2.4e-56;
Matches 226; Conservative 154; Mismatches 343; Indels 120; Gaps 17;

QY 214 IYTERERERE-----KSEKRRPMA-----VVC---TPQATHSLOPHHWP 257
DB 41 ISMTVEKASREDLAILVKGCDGQAMDTTRVFQMGIRNVITNRATGEIAHSTIYA 100
QY 258 SVRESLSTWPKNEDFNQKFTQLLLQRPDRSQDPLVKRSMPDYVENRGHLIEIRD 317
DB 101 HLKEL-----THDCSKF-----NISIQNFQDEYDHL---EVL 132
QY 318 FPGGLDTQEPRIYVILQGAAGIKSTIAROVKAWGRGQYGRFOHVFYFSGRELAQSV 377
DB 133 LVNNGTENNPKNVVLQGVAGIGKITILLKMLTWSEGLVFQNKFSYIFFCGHQDVQLO 192

```



```

QY 378 VSLAEIKDGTATPAPIRQLISPERLFTLDGVDGEMWVLOEBSSELCHMSQPOPAD 437
Db 193 ASLADLISREWPSSAPMEELISQPEKLLFTIDSEGMENWVTOODSOLCYNCKEOPVN 252
QY 438 ALLGSLIGKTLIPASFLITARTTALONLISLEQARAVEVIGFSESSEKREYFRFYFDE 497
Db 253 VLSLSLRKLLPSSSLISTSCETFDKLMIEWTNRITITGKRENNINCHSLRQDR 312
QY 498 ROAIRARLVKSNKELMALCLVPVWSWLAQTCLMOQMR-EKLTLSKTTTTLCLHYLA 556
Db 313 NIAOEASFLIRENQLFTVCOAPVYCWATCLNKEISGKDPVISCRTTSLYTHILN 372
QY 557 QALAQOLPGQ-----LRDLSLAESIGMOKTLFSPDDLRKXGLDCAIISTLXKML 610
Db 373 LFIHANONENNSBDDLDNLCFLAVEGMDVDSVFNEBARRNGIMDSIPTLLDIGIL 432
QY 611 -QEHPIPLSVSFHLCFOEFPAAMSYYLEDEKGRKHSNCIID-----LEKTEAVGIIH 663
Db 433 EQSRESENSYIFLHPSVQEFCAAFYLLHSEMDSCQGVFIETFTFLTKIKKQWVFL 492
QY 664 GLFGASTTRFLGLISDEGEREMENIFHCRLSQ--GRNLMQWVPSLOLLQP-----HSL 716
Db 493 GC-----FFGGLHETEQEKLEAFGYNHLSKELRQLFLMLRLDTHLPDYKXINTM 545
QY 717 ESHLCYETRNKTLTQVMAHFEEMGMCVETDMLVCTPCIKF-----SRHYKK 766
Db 546 KFFCCLTEMBEEVVOGAMNCREQIDVVVKYSDFIYAAYCLSHGSAITDPSISAQVNLN 605
QY 767 LQLEIGHQHSWTSPVTWVLFVWVWV-----TDAYWQILFVSULK-----805
Db 606 BELGQRKLLIMWQICSVFIRANDIKTIRIEDTIFNEPVKIKYTSYLKSSCTIKTLVA 665
QY 806 -----VTRNKEILDLSGNSLSHSAVSLCKTLRRPCLLETTLRLACGL 849
Db 666 YNVFSLCDKRLFLBILIOSYNLEBYLGTFLSHSDVENLCDILNQAEINRIIDLANCSL 725
QY 850 TAECKOLARGLRANQUTLELDLSPVLTDAAGHLQORLOPSCCKYQRLQVSCGLTSD 909
Db 726 CEHSWDYLSIDVLRNKSILRYINISYNNLKGBGLALCRALLTPSALSLSLEACQITGA 785
QY 910 CCODLASVLSAPSLKEHLDQNNLDVGVVLLCEGRHPACKLIRLGLDQTLSDEMRQ 969
Db 786 CCKDLASTFRRYKLRINLAKNSLGSGLFVLCMAKMDQCTIYELKLRMADE-DSDSQ 844
QY 970 ELRALBOEKPOL 982
Db 845 EFLSEMERNKIL 857

```

RESULT 7

Q91XTO PRELIMINARY; PRT; 846 AA.

AC Q91XTO; 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Similar to NALP2 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Strausberg R.;

RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC039269; AAH39269.1; -

SS SEQUENCE 846 AA; 96368 MW; 46BBE245503977 CRC64;

Query Match 11.2%; Score 843; DB 4; Length 846;

Best Local Similarity 30.0%; Pred. No. 1.8e-54;

Matches 255; Conservative 154; Mismatches 328; Indels 114; Gaps 28;

```

QY 148 WREISASLLQALPSSPDHSPSOESPNAPTSTAIVLGSWSPPOPSLAPEQEPGTOWP 207
Db 67 WVEWASLQVEEK-----HMDLSERAKDEVREALASFNKRKPLSIGITRKEP-----P 117
QY 208 LDETSGLIYTIEREEREREKSEKGRPPMAAVVGPPTQHTSLQPH-----HHMPBSVRES 262
Db 118 LD-----VDEMLERFTEKQAFETYGAVTCLAKEVFKGKPKDKNRCRYILKTKFRE- 170
QY 263 LCGTWPKNEDFN-----OKFTOLLQLQRPHPRSODPLVKSQMPDYVENRGLIEIRDLFG 319
Db 171 MWKSPDSDSKVQWMAHYKMLIFPSNP-----RVLPQ 203
QY 320 PGLDQTPRIYVILQAGAGTSTARQVKEAWRGQLYGDRFQGVFFPSCHLAQSKVYS 379
Db 204 PFSYT-----VVIYGPGLTKTLTAQMLMDMEDNLI-KHFKYAFLYSGSELSRLPCS 257
QY 380 LAELIGKGTATPAPIRQLISPERLFTLDGVDGEMWVLOEBSSELCHMSQPOPAD 437
Db 258 FAEIVFDMPELDQDDIPIHIAQARKILFVIDGFDELGAAPQALIE---DICGMEKKKP 313
QY 436 ADALLGSLIGKTLIPASFLITARTTALONLISLEQARAVEVIGFSESSEKREYFRFYFT 495
Db 314 VPVLGSLNRVWMLPKALLVTRPRALDRILAEPYIYIRVSGFLEEDRRAYFLAHFG 373
QY 496 DERQAITARLVKSNKELMALCLVPVWSWLAQTCLMOQMRKEXLTLSTTTTLCIHYL 555
Db 374 DEQAMRAFAELMSNALFOLGSAFAVCWIVCTTLKLQMEKGEQPVPTCLTRTQLFLFL 433
QY 556 AQAL-QAQLPGPOLRDCSLAAGSIWOKTLFSPDDLRKXGLDCAIISTLXKMLLEHP 614
Db 434 CSRPGQAQRLGALRTISLIAQGLMAQTSVLRHEDERGVQESDRLFLDGLDILQDR 493
QY 615 IPLS-YSFHLCFOEFPAAMSYYL-EDKGRGKHSNCIIDLEKTEAVGIIHGFCASTT 671
Db 494 VSRGCVSFHLSFOQFLTALFYTLKEEEDRQHTWIDQVXLIS-GVERLRNPDLI 551
QY 672 R---FLIGLISDEGEREMENIFHCRLSQGNLMQWVSL-QLLQ-----PHSL-----716
Db 552 QAGYVSGFLNERRAKELTATFCRMS-----FDIKQELRCDISCKGHSYVTDL 602
QY 717 -ESHLCYETRNKTLTQVMAHFEEMGMCVETDMLVCTPCIKFESHYKK--LOLI-EG 772
Db 603 QELIGCLYESQEBELVEXVNAQFKREISLHINA-VDVVPSFCVKGCHNLQMSIQVKEN 661
QY 773 ROHSTWSPVTWVLFVWVWV-----WVPTDAYWQILFVSULKYTRNLIKELDLSGNSLSHSAV 826
Db 662 LPENVTASSEDVAVERSGDDQHLP-----FWTDLCSIPGSKNDLMGAIINDSFLASIVR 717
QY 827 SLCKTLRRPCLLETTLRLACCGITABDC-KDLAFGLRANQUTLELDLSPVLTDAAGHL 885
Db 718 ILCEQIASDTCILQ--RVVFKNISPADAHRNLCIALRGHKTIVYTLTLOGNDODDM-FPAL 774
QY 886 CORLROPSCKLQRLQVSCGLTSDCCODLASVLSAPSLKEHLDQNNLDVGVVLLCEG 945
Db 775 CAVLRHPECNLRKLVGLVSCATITQWADUSLALVNOSSLTCVNLSDNELDBEAGKLLYTT 834
QY 946 LRHPACKLIRL 956
Db 835 LRHPKCTQLRL 845

```

RESULT 8

Q95L27 PRELIMINARY; PRT; 713 AA.

AC Q95L27; 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Hypothetical 80.3 kDa protein.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Macaca.

OX NCBI_TaxId=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 Terao K., Sugano S.;
 RT Isolation of novel full-length cDNA clones from macaque testis cDNA
 RT libraries.
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB071044; BAB6437.1; -
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR PROSITE; PS05053; LRR_R1; 2.
 DR PROSITE; PS0837; NACHT; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 713 AA; 80258 MW; 6F214C9B773F54DC CRC64;

Query Match 10.8%; Score 815.5; DB 6; Length 713;
 Best Local Similarity 31.4%; Pred. No. 1.6e-52;
 Matches 228; Conservative 114; Mismatches 248; Indels 135; Gaps 18;

QY 383 LIGKDGATPAPRIQILSRPERLLFLIDGVE---PGWVLQPPSELCLHNSQOPADA 438
 DB 1 MYFRDMPRLDDIPIHLAQOKILFVIDGFDELGAPFALIQ---DIGDMEQQKPVV 56
 QY 439 LIGSLIGKTLIPASFLITARTALONLIPSLQARWEVYLGFSSESKREYFRYTDER 498
 DB 57 LIGSLIKRKLPRATLLVTRPRALDRFLAQPIYIRVEGLEEDRRAYFLRHGDD 116
 QY 499 QAIRAFRLVSKNELMALCLVPMVSWLACTCLMQKKREKLLTSTKTTTLCIHL-AQ 557
 DB 117 QAMRABELMRKNALPOLGSAFVNCIVCTTLKQNEKEDRPCTLTGFLRLRCQ 176
 QY 558 ALQAPQLGRLDLSLAAGIWOXKTLSPDDLKRGHLDGAIISTFKMGLQEHPIPL 617
 DB 177 FPGAGLRGRLSLLAQSLWAQMSYVHGEDLESAGVQESDLRLFLDGDILRQGVAK 236
 QY 618 S-YSFHLCPQEPFAMSYVL--EDEKRGKHSNCIIDEKTL---EANGHGLFGASTT 671
 DB 237 GCASFHLSPQQLTALFYALEKEEBEDRGHAMDIGDVQKLRSREERLKNPDLIQAG-- 294
 QY 672 RFLGLLSDEGEREMENIFHCRLSGQRNLMQWPSL-QLLQ-----PHSLRS 718
 DB 295 RFLFGLANERKVELLATRGMRMS-----PEIKQELLRCDSVRKNGHPTAALREL 345
 QY 719 LHCYLTFRNKTLTQVVAHEENKMCVETDMELLVCTFCIKFSRAV-----KGLQ--- 769
 DB 346 LCLYSESQDELVEVMAQFKEISLHLNA--VDIAPSSFCFKHCQNIFFPADACNLLCLALR 404
 QY 770 -----IEGRQHSSTSPVTVL-----FRWV-----PYTDAYWQILPSVLKVTN 809
 DB 405 GHKTVTLTLQGTDOQDMPLALCEVLRHPECNLRVGLVSCSATTOQWMDLSLALRANS 464
 QY 810 LKELDLSGNSLSHSAVSLCKTLRPRCLLETRLAGCGLTAECDKLAFLGLANGLTNE 869
 DB 465 LMCVNSIDNEHLDGAKLYTLTRHKPCFLQRLSLNCHTEANCDLAAVLVVSBLTH 524
 QY 870 LQLSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSELKELD 929
 DB 525 LCLAKNSLKDQGVKFLCEGLSPECKLQALVLMNCITSDGCCSLAKLQEKSLSLCLD 584
 QY 930 QQNNLDVGVRLLCGLRHPRCKL----- 953
 DB 585 GLNHIGVAGVLCALSKPLCNLRCLMWGCSIPPSCEDLCSALSCNOSTLTLQON 644
 QY 954 -----IRLGDOTTLSDEMROELRALOEKPOLLIPS----- 985
 DB 645 PLGSSGVKMLFKLTLPGLTGLTKIKID--FNDELHKLLEIENNPQILIDTEKHDPV 702
 QY 986 RRRPS 990
 DB 703 KRRPS 707

RESULT 9
 Q99MWO ID Q99MWO PRELIMINARY; PRT; 748 AA.
 AC Q99MWO;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ribonuclease/angiotensin inhibitor 2.
 GN RNH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=21175748; PubMed=11279525;
 RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
 RT "An abundance of X-linked genes expressed in spermatogonia."
 RL Nat. Genet. 27:422-426(2001).
 DR EMBL; AF285581; AAK31960.1; -
 DR HSSP; P10775; 2BNH.
 DR MGD; MGI:1890518; Rnh2.
 DR InterPro; IPR007091; LRR_RNinh.
 DR PROSITE; PS05053; LRR_R1; 1.
 DR PROSITE; PS0837; NACHT; 1.
 SQ SEQUENCE 748 AA; 84946 MW; 674A3B5E011DE8EC CRC64;

Query Match 9.2%; Score 689.5; DB 11; Length 748;
 Best Local Similarity 26.2%; Pred. No. 5.6e-43;
 Matches 197; Conservative 120; Mismatches 241; Indels 193; Gaps 14;

QY 419 LQPSSELCLHNSQOPADALIGSLIGKTLIPASFLITARTALONLIPSLQARWEV 478
 DB 5 MERSESLCDTCTEKQPRILSLSLRRKMLPSFLSATPTEKMGREACTVAKIV 64
 QY 479 LGFSSSRKREYRRFTTERQALRAFLVYSKXELMALCLVPMVSWLACTCLMQQK-KR 537
 DB 65 TGFNSNINMYRSLPQDKTKQELFSLVKENQQLFTVCQVVLCCMWATCLAKTEKER 124
 QY 538 EKLTLSTKTTTLCILH---YLAQALQ--APQLGRLDLSLAAGIWOXKTLSPDDL 591
 DB 125 DLVSQCRRTSLYTHHIFNLPIQSAQYPSKESQALQSLGSLAEGMWTDFVGEAL 184
 QY 592 RRGHLDGAILSTFKMGLIQ--HPIPLSYSPHLCPQEPFAMSYVLEDEKRGKHSNCI 650
 DB 185 RRGIMDSDIPTLIDVRLIEKSKSEKSYIFLHPSIQEVCALIFYL--KSHMHPQSD 241
 QY 651 IDLEKTLA-----YGIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQ--GRNL 700
 DB 242 V--KSIELIIFTLKRVKVMIFPGS--FTGLHSEBQKCLAFQHQLSQSLKRL 295
 QY 701 MQWVPSL-----QLLQPSLSLSLHCLYETRNKTLTQVVAHEENKMCVETDMELLVCTF 756
 DB 296 YQCLFETISNEELQOVGMKLFYCLLFEMDEDAFLAQMNCMEQINFPAKQVSDIYVAH 355
 QY 757 CIKFSRHVKQLQ-----LISGRHRSST--N----- 779
 DB 356 CLQHSSTLKKSLSTQNVLSBQHSYTEKLMCMHMSVLISSKDIYILQVKNNTINE 415
 QY 780 ----- 779
 DB 416 TSLVLYSHLMPSECTLALVYNNVTPFLCDNRLFELLIONCLQHLDLNLTLSHQDVLT 475
 QY 780 -----SPTMVVLFRRV----- 790
 DB 476 LCDVLSQEBENIEKLMVAACNLSPPDDCKVFAVLISSKKLKHILANSSNNLDKGISSLSKA 535
 QY 791 -----PYTDAYWQILPSVLKVTRLKELDLSGNSLSHSAVSLCKTLR 834
 DB 536 LCHPDCVLKNLVVNCSSLEQCMVDLSEVLRNKKTLNHLDISNDLKQEGLVLCRAVSL 595
 QY 835 PRLLETTLRAGCGLTAECDKLAFLGLANGLTLELDSFNVLTDAGAKHLQRLRQPS 894

```

Db      596 PDSVLSLSRYGLTITSSGQDLAEVLARKQNNRLQVSNKKEEDAGVKLLCAIHRPNC 655
QY      895 KQRLQVLSGGLTSDCCODIASVLSASPISKELDLQNNIDVGVRLLCGRLHPACKLT 954
Db      656 HLENIGLEACALGACCCEDLASAFTHCKTLMGINTQENALDHSGLIYLFALPKQOCTH 715
QY      955 RLGDOTTLSDEMRQELRALEQEKPOLIFS 985
Db      716 VLGRIITDFEKTQELMAEBEKPHILSTLS 746

RESULT 10
Q8CGMS PRELIMINARY; PRT; 825 AA.
ID Q8CGMS
AC Q8CGMS;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Weakly similar to PANI.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054264; BAC35710.1;
SQ SEQUENCE 825 AA; 94021 MW; 11D71DEABEAFCA31 CRC64;

Query Match 9.0%; Score 679.5; DB 11; Length 825;
Best Local Similarity 26.0%; Pred. No. 3.7e-42;
Matches 212; Conservative 150; Mismatches 313; Indels 141; Gaps 24;

QY 208 LDETSGLTYTEIREREKSEKRPMAAVVGTTPQAHTSLQPHHWPBVSRESLCTW 267
Db 76 LEBINRVDLAEVLVKQEB-----ATLKVPBEEKVS-----KRPBDSG-----TLTF 117
QY 268 PMKNEBPNOKFTQULLQRPHPRSQDPLVK-----SWPD-----YVENRGHLIE 313
Db 118 PM-----NFVQ--GAKRPEDQKQEKWKTRYAKWKQNFMPKCKEIIYVVTESYKTLA 168
QY 314 IRDLFGPGLDTPRPIYILQAGAGIGKSTLARQVKEAMGQLYGDRFGVVFQSCRELA 373
Db 169 ---LCNPKLSTPRRAHAIVLHGPPSGKTYAKQKQMLMEWSESK--QAQIFSCAFYISCREVN 224
QY 374 QSKVSLAEILGKDGATAPAPIRQLISPERLFIIDGVDEPGVLOEPSSSELCIHWSCP 433
Db 225 NTRKPCFPAHLISMDNDSWRCVIRDLIGKEFLFYVDGPFELFPAGALIRDLQGMWTV 284
QY 434 QRPDALIGSLGKTIIPASFLITARTTALQNTIPSLQARWVEVLGSSSKKEYFYKY 493
Db 285 KPRVVLIGSLKRRKMAPHATLLVTRFOSIHOIFVMMDDPLVETLGFLEQEOEYFOKY 344
QY 494 FTDER-----QAIRAFLVSKNELMALCLVPSVSWLACTCLMQQKRRKKTLLTS 544
Db 345 FEBEBEGBEKGKAKIRALKVRCNADLYQMSLTPACGIFCLCELRKKGSDLSLTIC 404
QY 545 KITTTTCLHLYLAQALQAPLGPQRLD-----LCSLAEGIWOKKTLFSPDDLRRHGL 596
Db 405 QYVTSWFLNLFCEVFSSETCEDHNEBFQILFKKICLIANSLSLEQVPILCEEDF----- 459
QY 597 DGAISTFLMGLIOEHPILSY-----STIHLCFQOFPAASVYLE---DEK 641
Db 460 -----LTLKLNNHNPWCRIHLFPDSSSTHCLSPICGICQQLAAIIFVDELGQESK 513
QY 642 GRGKHS-NCIIDEKTLVAYGIHQLFGASTTRFLLGLSDGEBREMNIFHCLS----- 695

```

```

Db      514 GYSKVIQWMLSEALNKPRLDLSGLL-----PVRGLMNETRLOELKTTTGCISTEYVR 568
QY      696 -----QGRNLQWVPSLQLLQPHSLESNCLYETRNKTFILTOVNAHEEMQCVETDM 749
Db      569 KFLCESGEN-----KELLIMNMQEILSCYSEQEBGFVKEAVLFPEDISLHKTST 621
QY      750 ELLVCTPCLKFSNHYKQLQ-IF-----GROHSTWSPTVVVLRWVPVTD 794
Db      622 DLTHASFCLKNQSLQITMSLKVKAVPENVALSESTARQORSDBQRM----- 671
QY      795 AYWQILFVYLKVRNLTKELDLSGNSLSHSAVSKLCTKLRPRCLTETRLAGCGTLAEDC 854
Db      672 TFWTDCDTPNSKULVFYLDIHESFLNSSLIELCEKLSASCCLOKVILK--NISPDDA 729
QY      855 -KDLAFLRANQTLTLELDLSFNYLTDGAKNLCQRLRQPSCKLQRLQVLSGGLTSDCCOD 913
Db      730 YEKLCILFNGYKTIISHLILOGNL-DSMHNSICEVLKNPACNLKFLSLGSCSTAAQWMD 788
QY      914 LASVLSASPSIKELDQNNNDVGVRLLCGRLHP 949
Db      789 FEPVLKVNQSLIFLDTLDRSLDKSAKLLCNITWKEP 824

RESULT 11
Q9EBG7 PRELIMINARY; PRT; 657 AA.
ID Q9EBG7
AC Q9EBG7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 23, Last annotation update)
DE RNI-like protein.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RA MEDLINE=21310002; PubMed=11416212;
RA Lane R.P., Cutforth T., Young J., Athanasiou M., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT loci."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395(2001).
DR EMBL; AF321233; AAG45188.1;
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR PROSITE; PS50503; LRR_RT_1.
DR PROSITE; PS50837; NACHT_1.
SQ SEQUENCE 657 AA; 75410 MW; F4BED9E4BA19AAF3 CRC64;

Query Match 9.0%; Score 676; DB 11; Length 657;
Best Local Similarity 27.1%; Pred. No. 4.8e-42;
Matches 194; Conservative 95; Mismatches 220; Indels 206; Gaps 11;

QY 314 IRDLFGPGLDTPRPIYILQAGAGIGKSTLARQVKEAMGQLYGDRFGVVFQSCREL 372
Db 78 LQHFIPEDIQTSEAPQTVVLQAGAGIGKTTLLKKALEWADGNTLY-QQFTHVLYLNGKEI 136
QY 373 AQSIVSLAEILGKDGATAPAPIRQLISPERLFIIDGVDEPGVLOEPSSSELCIHWSC 432
Db 137 SQVKEKSPHQLSKMWPSSGEPTEQ-----QVKEGGRVER----- 172
QY 433 POPADALIGSLGKTIIPASFLITARTTALQNTIPSLQARWVEVLGSSSKKEYFYKY 492
Db 173 ----- 172
QY 493 YFTDERQAIRAFLVSKNELMALCLVPSVSWLACTCLMQQKRRKKTLLTSKTTTLLC 552
Db 173 -----TCKTSTALFT 182
QY 553 HYLAQALQAPLG-----PQRLDCLSLAAGSIWOKKTLFSPDDLRRKHGLDGAISTFL 605

```

```

Db      183 YIICSLFPRIPVGCVTLLPNETLRSLSCKAAVEGIMTKKAVLYQONRKHELTREDILLRL 242
Qy      606 KNGIIOEH-PIPLSYSPHLCFOEFPAAMS-YLV-EDKGRGKSNCIIDLEKTLBAVGHI 663
Db      243 DAKVILQODTEYENCYMFHLHVQEFPAALFYLLREMLEEDYSEPFENLYLLSENNH 302
Qy      664 GLFGASTTFLGLLSDGEGREMENTFHCKLSQ--GRNLMQWVPSIQ---LLQPHSL 717
Db      303 DPHLEOMKCFELFGLLNKDRVRQLEEFNLTISMEVEEELACLEGLEKSDSSLSQRFOD 362
Qy      718 SLHCLYETRNKTEFLTOVMAHFEEMGMCVETDMELLCTFCIKFSRMYKQLOLEGRQHS 777
Db      363 LHCIYETQDERITQALMYFOKIYRVDBEPOLRYSFCLKCHTLKTRMLTARADLKH 422
Qy      778 TWSPTNVLFRWVPVTDAYWQIIFSVLKYTRNLKELD----- 815
Db      423 MLDTAEMCEGAAYVYHWDLFVSYLHTMESLIEMDLVESRLDESIMKLINELSHPKC 482
Qy      816 -----SGNSLSH-----SAVKSCLTKLRPRCLLETL 842
Db      483 KLQKLIFRSVDPLNGODFTFLASNKKVTHLDKETDLGVNGIKTLCBALCKCKGCTLRVL 542
Qy      843 RLACGGLTAEDECCKDLAFGLRANOTLTELDLSFNVLTDAAGKHLQORLQPSCKLQRLQV 902
Db      543 RLASCDLVARCKSLSNALQTRNSLVFNLISLNSNDGVKSLCEVLEMPNSLSERLALM 602
Qy      903 SCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLLCEGLRHPACKLRLG 957
Db      603 SCVLTSKACODLASVLYVNSNLSLIDGHIILDDAGLNLCDALRNPCHQVRLG 657

```

RESULT 12

```

Q8CCN1 PRELIMINARY; PRT; 673 AA.
AC 08CCN1;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE PAD and NACHT containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=2354683; PubMed=1246681;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573 (2002)."
DR EMBL; AK032446; BAC27872.1; -.
SQ SEQUENCE 673 AA; 76367 MW; EE773C9S2BEC7054 CRC64;

```

Query Match 7.6%; Score 573; DB 11; Length 673;
 Best Local Similarity 33.9%; Pred. No. 2.9e-34;
 Matches 152; Conservative 81; Mismatches 154; Indels 62; Gaps 14;

```

Qy      312 IEIRDFGGLDT-QEPRVYILOGAAGIGKSTLARQVKAMRGOLYGRFQHVFEFSGR 370
Db      146 VDEETIFAPBAESYSTPPIVWQSGAGTGKTVLVKGLVDQMSGKLYPGQFDVIFYVSCR 205
Qy      371 ELAOSKVSIAELI-----GKDGATPAPTRQILSRPERLLFINDVDEBGWVLOBPSESL 426
Db      206 EVLLPKCDLPNLICGCCGD---QAPVTEILRQGRLLFLIDGVE---LQKSSR-- 255
Qy      427 CLHMSQPADALLSLGKTIIPKASFLITATTLQNLIPLEBARWVYLGSESSR 486
Db      256 -----AECVLHILMRREVP-CLSLITTRPALQSLLEPMIGERRHYLVYGFSEBR 305
Qy      487 KEYFYRYFTDERQAIARLVKSNKELMALCLVPWVSWLACTCLMQMKKELTLTSKT 546

```

```

Db      306 EYFSSCFDQKQALAEFYQNNALVYKACVPGICWVVCWMLKKKARQGEVSETPSN 365
Qy      547 TTTLCIHLAQLAQAPLPQ-----LRDLSLAAGIWOQKTLFSPDDLKRGGLDG 598
Db      366 SDIDIFAYVSTLPDNGDSSSELTRHKVLSKSLAAGMHQHLFPBEVLRKHGLDG 425
Qy      599 AIISTPLK-----MGIIOEHPIPLSYSPHLCFOEFPAAMS-YLV-EDKGRGKSN 648
Db      426 PELTAFPLNCIDYRAGIGIK-----FYSFRHISFQEFFYAMSFLVKEDQSGGEATHKE 479
Qy      649 C-IIDLEKTLBAVYGHIGLFGASTTRPFLGLSDGEGREMENTFHCRU--SGRNLMQWV 704
Db      480 VAKLVDPENHEEV-----TSLQFLPDMKTKGTSLGLKFCRLAPSVRODLKHPK 531
Qy      705 PSLQILLQPHSLSLHCLYETRNKTEFLQ 733
Db      532 EQIEAIKVRSDWDLFSLYDSKIK-LNQ 559

```

RESULT 13

```

Q8IMF5 PRELIMINARY; PRT; 953 AA.
AC 08IMF5;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Caspase recruitment domain family, member 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RT Submitted (NOV-2002) to the EMBL/GenBank/DBSj databases.
DR EMBL; BC040339; AAH40339.1; -.
SQ SEQUENCE 953 AA; 107671 MW; 0A9DF167BE87E21A CRC64;

```

Query Match 5.2%; Score 393.5; DB 4; Length 953;
 Best Local Similarity 24.0%; Pred. No. 1.8e-20;
 Matches 244; Conservative 131; Mismatches 336; Indels 209; Gaps 38;

```

Qy      237 VGTTPQAHTSLQPHHWPBESVRESLSCTWPKNEDFNQKFTQLLLQRPHPQODPLY 296
Db      120 VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAKK--ELLLEIYMDTIMELY 167
Qy      297 KRWEDYVBEKRGHLEIRDLFG--PGLDTQEPRIYILOGAAGIGKSTLARQVKAMRG 354
Db      168 -----GFSNESLGSINSIACLDHTTGILNEQGTIFILGAGVSKMLQRLQSLMTG 222
Qy      355 QV-YGDRFOHVYFSCRELA--OSKVSIAELICKD--GATPAPTRQILSR-PERLL 406
Db      223 RLDAQVKF--PFFHCNMFSCSKESDRCLQDLDFKHVCYPERDDEVPFALFRPHVAL 280
Qy      407 FLIDGVDPRGWVLOBPSESLCH-----WSQPADALLSLGKTIIPKASFLITA 458
Db      281 FTFPDGLD-----ELHSDLDLSRVPDSSCEPBNAPL-VILLANLISGLTKGASKLITLA 332
Qy      459 PRTALQNLIPSLAQKRWYEVIGFSSSRKEYRYRFTDERQAIARFLVKSNEKELMALCL 518
Db      333 RTGI--EVPQPLKRVKLVNGFSPSHLRAVARMPFERALQDRLLSQLEANPNLCSLS 389
Qy      519 VFWVSWLACTCLMQ-----QMKRKEKTLTYSKTTTTL 550
Db      390 VFLFCMIIFRCQHRAAFBESQPLPDCMTLLTVDFLLVTEVHLNARMQPSLSVQNT--- 446
Qy      551 CLHYLAQLAQAPLPQLD-LCSL--AASGIWQKTLFSPDDLKRGGLDGAIISTPLK 606
Db      447 --HSFVETLHAG-----RDTLCSLGVYAHNGMEKSLVFVQEEVQASGLQ--ERDMQ 494
Qy      607 MGIQENP-----IPLSYSPHLCFOEFPAAMS-YLVDEKGR-----GK 645

```

```

Db      495 LGFRLALPELGGDQGSYEFPHILLOAFAPFLVDDRGVGTQELARFROEMPPAGAA 554
Qy      646 HSNCT-----IDLEKLEAYGHLFGASTTRFLGLSDGE-----683
Db      555 TTSCYPPFLPFOCLGSGPAREDLFKNKD---HFQF---TNLFGLSLXAKOKLRL 607
Qy      664 -----REMENTHCRSLSGRNLMOQWPSL-----QLLQPSLSLEHCLYETRKTFF 730
Db      608 VPAAALRRKALMAHLPSSLRGYKSLPRVQVESFNQVQMPFFIMWLRCIYEQ----663
Qy      731 LTVMAHFEEMGMCVETDMLVCTFC-----IKFSRH-VKKLQL-----TEGR 773
Db      664 -SQKVGGLAANGICANY-LKLYCNACGADCSALSFVHHFPKRLALDDNNMLNDVGR 721
Qy      774 QHSTWEPYVLFRRVPTDAYVQILFSLVKTNRNKELDLSGNSLSHSAVKSLCKTLR 833
Db      722 ELQCFRSLTVLRISVNOITDGVKVLSEELTKYKIVLYGLYNNQITLDVGARYTKLD 781
Qy      834 RPRCLLTLRLAGGLTAEDCKDLAFGLRANQITLEDLSFNVLTDGAKHLQRLR-QP 892
Db      782 ECKGLTH-LKIKGNKITSBGKYLALAVKNSKSISEVGMMGNQVGDGAKAFALRNHP 840
Qy      893 SCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGYRLCEGLR-HPAC 951
Db      841 S--LTTLSLNSNGISTEGKSLARALQONTSLLETIMLTQNELNDVASELSAEMLKNOVL 898
Qy      952 KLILGLDQTT-----LSDEMR-----QELRALQOK 978
Db      899 KHLMLIONQITAKGTAQALADALQNTGITCINGNLIKPEAKYEDK 948

RESULT 14
Q8K320 PRELIMINARY; PRT; 1020 AA.
AC Q8K320;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE CARD15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Ivanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15
RT gene."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF520774; AAM76073.1; -.
DR MGI; MGI:2429397; Card15.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR001130; Tact DNase.
DR PROSITE; PS50209; CARD; 2.
DR PROSITE; PS50503; LRR_RI; 1.
DR PROSITE; PS50837; NACHT; 1.
DR PROSITE; PS01137; TATD; 1.
SQ SEQUENCE 1020 AA; 113561 MW; 25504905ECF70FBB CRC64;

Query Match 5.2%; Score 392; DB 11; Length 1020;
Best Local Similarity 25.2%; Pred. No. 2.5e-20;
Matches 186; Conservative 122; Mismatches 316; Indels 114; Gaps 28;

Qy      311 LIEIRDLGPPGLDTQEPRIYILGGAIGKSTLAROYKEAMGSGQYGRFQHVFFSC 369
Db      255 ILGHEDLFDTHGNLRADATILVVGAGSGKSTLQRLHLMLMGSRSF-QEFLFIFPFS 313

```

```

Qy      370 REL-AQSKVSLAELICK-----DOTATPAPITROILSRPEELLFTLDGVDERGWLQEPSS 424
Db      314 RQLOCAKPPSLKTLLEHHCWPDVADVDVQFLDDPRDLTFEDGLDEKFEETRER 373
Qy      425 ELCLHWS--OPPADALGSLGKTLIPKASFLITATTTLOMLIPSLQARWVEVLGFS 482
Db      374 ---HCSPIPTSGVTLTFRLQGNLILKNAOKVLTSPDVAVALRRFVTE-LQLGFS 428
Qy      483 ESRKEFYFRFTDERQALTAFLVSNKELMALCLVFWSWLACTGLMOQMKREKLT 542
Db      429 EBGQVLRKHREBPVADRLIQLTATSLHGHCHLPVPSMWSRCHRELLQNRGPT 488
Qy      543 TSKTTTLCL-HYLAQLQAP-----LGP-----QLNDLCSLAEGTQKTLF 586
Db      489 TSDMYILLIQHFL---LHASPPDSPLGLGPGLLGSLSTLHLGLALRGMLSCYVF 545
Qy      587 SPDDLKRGHGLDGAIISTFLKMGIL--QEHPIPLS---YSFTHLCFOEFPAAMSYLEDEK 641
Db      546 SAQQLQAQVDSDDIS---LGFVLRAGSSVPSGKAPLEFLHITTFQCFPAFLVAVADT 601
Qy      642 -----GRKHSNCT-IDLEKLEAYGI---HGLFGASTRFL 675
Db      602 SVASLKLFPSCRLGSSILGRLLPNLCIQSRVKGSEALLQKABPHN--QITTAFLA 659
Qy      676 GLSDGEREMENIFHCRLSGRNLMOQWPSLQLLOPHSL-----ESLH 720
Db      660 GLLS---QQRDLIAQVSE-RVLLQROARASCL-AHLREHFHSIPPAVPGETYSMH 714
Qy      721 C-----LYETRNKTFLOVMAHFE---EMGCVETDMLVCTCFIKFSRAVKK 766
Db      715 AMPGFWLIRSLYEMGEQALQEAIVRRLDIGHLKTFCRGPACALAVLOHLQRPVA 774
Qy      767 LQI-----IEGRHRSWSPMTVVLFRRVPTDAYVQILFSLVKTNRNKELDLSGNS 819
Db      775 LQDYNVSVDGVGEQLEPCLGCTALYLRDNNISDRGAPFLVCALRCEQLKALFPNNK 834
Qy      820 LSHSAVSLCKTLRRRCLETLRIAGCGLTAEDCKDLAAGLRANQITLEDLSFNVLTD 879
Db      835 LTPACACSMKLAHKONFL-SLRVGNHHTTAAGAEVLQGLSNSTLKLFGWNGSVGD 893
Qy      880 AGAKHLQRLQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGV 939
Db      894 KGYQALAEVADNQ-NLXWLSLVGNNGISMGABALMLMEKNLSLELCBENHICDEGV 952
Qy      940 RLICEGL-RHPACKLIRL 956
Db      953 YSLAEGIKRNSLTKFLKL 970

RESULT 15
Q9BY26 PRELIMINARY; PRT; 287 AA.
AC Q9BY26;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Leucine-rich-repeat protein RNO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=;
RA Shami P.J., Kanai N., Wang L.Y., Vreeke T.M., Parker C.J.;
RT "Identification and characterization of a novel gene that is
RT upregulated in leukemia cells by nitric oxide."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF31021; AAK14942.1; -.
DR HSP; P13489; I44Y.
DR InterPro; IPR007091; LRR_RNinh.
DR PROSITE; PS50503; LRR_RI; 2.
SQ SEQUENCE 287 AA; 31765 MW; BD3816C3255B2F9E CRC64;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:11:31 (Search time 28.1576 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title: US-09-996-617-2

Sequence: 1 MAGAGWRLACYLEPLKKEE.....HLIMELMKSKKKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	388.5	5.2	953	4 US-09-099-041A-8	Sequence 8, Appl1
2	388.5	5.2	953	4 US-09-245-281-8	Sequence 8, Appl1
3	388.5	5.2	953	4 US-09-207-3598-8	Sequence 8, Appl1
4	388.5	5.2	953	4 US-09-340-620A-8	Sequence 8, Appl1
5	363.5	4.8	456	2 US-08-910-731-8	Sequence 8, Appl1
6	361.5	4.8	456	2 US-08-910-731-4	Sequence 4, Appl1
7	361.5	4.8	456	2 US-08-795-395-4	Sequence 4, Appl1
8	357	4.7	461	2 US-08-910-731-6	Sequence 6, Appl1
9	354.5	4.7	461	2 US-08-910-731-2	Sequence 2, Appl1
10	354.5	4.7	461	2 US-08-795-395-2	Sequence 2, Appl1
11	348	4.6	953	4 US-09-245-281-43	Sequence 43, Appl1
12	348	4.6	953	4 US-09-207-3598-43	Sequence 43, Appl1
13	348	4.6	953	4 US-09-340-620A-43	Sequence 43, Appl1
14	345	4.6	1130	2 US-08-519-547A-6	Sequence 6, Appl1
15	326.5	4.3	966	4 US-09-207-3598-47	Sequence 47, Appl1
16	326.5	4.3	966	4 US-09-340-620A-47	Sequence 47, Appl1
17	250.5	3.3	135	4 US-09-340-620A-49	Sequence 49, Appl1
18	218.5	2.9	430	4 US-09-099-041A-26	Sequence 26, Appl1
19	218.5	2.9	430	4 US-09-245-281-26	Sequence 26, Appl1
20	218.5	2.9	430	4 US-09-207-3598-26	Sequence 26, Appl1
21	218.5	2.9	430	4 US-09-340-620A-26	Sequence 26, Appl1
22	216	2.9	71	4 US-09-340-620A-58	Sequence 58, Appl1
23	209	2.8	70	4 US-09-340-620A-57	Sequence 57, Appl1
24	209	2.8	70	4 US-09-340-620A-66	Sequence 66, Appl1
25	198	2.6	1151	3 US-08-836-134-23	Sequence 23, Appl1
26	198	2.6	1151	3 US-09-493-784-23	Sequence 23, Appl1
27	198	2.6	1232	3 US-08-836-134-2	Sequence 2, Appl1

28	198	2.6	1232	4 US-09-493-784-2	Sequence 2, Appl1
29	179.5	2.4	483	4 US-09-304-615-154	Sequence 154, Appl1
30	175.5	2.3	200	4 US-09-099-041A-11	Sequence 11, Appl1
31	175.5	2.3	200	4 US-09-245-281-11	Sequence 11, Appl1
32	175.5	2.3	200	4 US-09-207-3598-11	Sequence 11, Appl1
33	175.5	2.3	200	4 US-09-340-620A-11	Sequence 11, Appl1
34	165	2.2	1466	4 US-09-252-991A-30085	Sequence 30085, A
35	147.5	2.0	793	3 US-09-012-710-10	Sequence 10, Appl1
36	147.5	2.0	793	4 US-09-556-273-10	Sequence 10, Appl1
37	147	2.0	2482	1 US-08-328-254-6	Sequence 6, Appl1
38	143.5	1.9	4302	4 US-09-052-469-8	Sequence 8, Appl1
39	143.5	1.9	4302	4 US-08-422-582-8	Sequence 8, Appl1
40	143.5	1.9	4339	4 US-09-052-469-6	Sequence 6, Appl1
41	143.5	1.9	4339	4 US-08-422-582-6	Sequence 6, Appl1
42	142.5	1.9	4302	3 US-08-658-136-5	Sequence 5, Appl1
43	142.5	1.9	4303	2 US-08-460-751-2	Sequence 2, Appl1
44	138.5	1.8	794	1 US-08-393-333-2	Sequence 2, Appl1
45	138.5	1.8	794	3 US-09-087-465-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1
US-09-099-041A-8
; Sequence 8, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099, 041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019, 942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Pred. No. 3.7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;

QY	237	VVGTPQAHSLQPHHPMPVSRESLCTWPKNEDFNOKFTQLLRPHRSDDPLV	296
DB	120	VVNTDVSRYTQQLRHHLGRDS-KFVLC-----YAKK--EELLLEITMDTIELV	167
QY	297	KRWPDYVENRNGHLIIRDLPF--PGLDQEBRIYLOGAGIGKSTLARYKAWMGRC	354
DB	168	-----GFSNBSIGSLNSLACLDHTTGILNEGETIFILGDAGVGSMLLQRLQSLWATG	222
QY	355	QL-YGRPQHVFFYFSCRELA---QSKVSLAELIGD--GRATPPIROIISR-PERLL	406
DB	223	RLDAGYKF--FFHRCRMFSCFESDRLCIQDLFFGYCYPERDPPEVFAFLRPFVVAL	280
QY	407	FILDGDEPGWVLQESSSELCH-----NSQPPADALISLIGKTLPEASFILTA	458
DB	281	FTFDGMD-----ELHSDLDLSRVDDSSCPWEPANPL-VLLANLISGLIKLKGASKILTA	332
QY	459	RTTALQNLIPSLQARWVVLGFSSESSRKEYFYFTDROAIRARLYVSKNELMALCT	518
DB	333	RTGI---EVPRQRLRKVLLRGSPSHLRAVARMPEBALQRLLSQLEANPNLCSLGS	389
QY	519	VPPVNSLACTCLMQMKRKE-----KLITLSTKTTTLCLHY-----LAQALQAP	563
DB	390	VPLFCWIIIRRCQHPFAFEGSPQLPDCMTITLDELVEVLTAVLNMOSSLSQRTNRSP	449
QY	564	L-----GQLRD-LGSL---AAGIWQKTLFSPDRLRGHGDGALISTFLKKGILQENH	614


```

Db      450 VETLHAG---RDTLSLGGVAHKGMEKSLFVFQSEVQASGLQ---ERDMQJGFLRALP 502
Qy      615 -----IPLSYSFHLCPQEFPPAAMSYYLDEKGR-----GKHSNCI----- 650
Db      503 ELPGSGDQGSYEFPHLTLQAFPTAFVLVDDRVTQELLRFQEMWMPAGAATSCYPPF 562
Qy      651 -----IDLEKTLBAYGIGHLFGASTTRFLGLSDGE-----R 684
Db      563 LPFOCLQSGPARBDFKXKD---HFQF---TNLFLGCLSKAKQKLRHLVPAALRR 615
Qy      685 EMENTFHCRLSGGRNLMQWPSL-----QLLQPHSLBSLHCLYETRNKFTLQVMAHF 738
Db      616 KKKALMAHLPSSLRGYLKSLEPRVQESFNQVQAMPFTIMLRCTYEQ-----SQKVGOL 670
Qy      739 EEMGMCVETDMLLVCTFC-----IKFSRH--VKKLOL-----IEGRQHSWTSP 781
Db      671 AARGICANY-LKLTVCNACSDCASLSFVHLHFPKRLALDNNNNDVGAHELQCFGR 729
Qy      782 TMYVLFPMVPTDAYWQILFSLVLTNRNKELDLGSNSLSHSAVSKLCTLRPRCLLET 841
Db      730 LTVLRLSVNQITDGGVKVISEELTKYKIVTYGLVNNQITDVGARYVTILDECKGLTH- 788
Qy      842 LRLAGCGLTAEDCKOLAFGLRANQTLTELDSFNVLTDGAKHLQORLR-OPSCKLQRLQ 900
Db      789 LKLGKXKITSBEGKTLALAVKSKSISEVGMNGQVGDGAKAPALNHRHS--LTTLS 846
Qy      901 LVSCGLTSDCCDGLASVLSASPSLKELDLQNNLDDVGRLLCEGLR-HPACKLIRLGLD 959
Db      847 LASNGISTEGKSLARALQONTSLIETLWLTQNELNDEVAESLAEMLKVNQTLKHLMLION 906
Qy      960 QTT-----LSDBRM-----OELRALQOK 978
Db      907 QITAKGTALDALQSNLTGITEICLNGNLKPEAKVYEDEK 948

```

```

RESULT 2
US-09-245-281-8
; Sequence 8, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-245-281-8

```

```

Query Match      5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Freq. No. 3,7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
Qy      237 VVGTTPQATSLQPHHWPMSVRESLCTWPKNDFQKFTOLLILQRPHPRSODPHY 296
Db      120 VVNTDVSRYTQLRHLGRDS-KFTLC-----VAK--ELLLEIYMDITMELY 167
Qy      297 KSMWPDYVENNCHLIEIDLFQ--FGLDQEFRIVLQAGAGIGKSTLAAQVKEAWGRG 354
Db      168 -----GFSNESIGSLNSLACLDDHTTGILNEQGETFIILGDAGVGKSMILQRLQSLMATG 222

```

```

Qy      355 QI-YEDRFQHVYFESCRELA---QSKVSLAELIGKD---GTATPAPIRQILSR-PERIL 406
Db      223 RLDAGVKF--PFFHCFRMFSCFKESDRLOLDFKHVCYPRBDEEVAFLRPFHVAL 280
Qy      407 FLIDGVDDEGWLOQSSSELCH-----WSQOPADALLGSLGKTLLEASTLIRA 458
Db      281 FTFDGLD-----ELHSDLDLSRVDPDSCPEBPHPL-VLANLISGKLLGASGLTRA 332
Qy      459 RTTALQNLIPSLQARWVEVLGFSSESRKEYRYRYFTDERQAIRAFLVSKKELMALCL 518
Db      333 RTGI---EYPRQFLKKTYLIRGFSFSLRAYARMPERALQDRLLSQLEANPNLCSLTS 389
Qy      519 VPMWSWLACTCMQOMKRE-----KLITSYTTTTLCHY-----LAQALOAP 563
Db      390 VLPFCMIIFRCFQHPRAAFEGSPOLPDCMTLITDVLVTVBHLNRMPSSLVORNTSP 449
Qy      564 L-----GQPLRD-LQSL--AAEGIMOKTLESPPDLRKHGDGAIISFLWIGLIEHP 614
Db      450 VETLHAG---RDTLSLGGVAHKGMEKSLFVFQSEVQASGLQ---ERDMQJGFLRALP 502
Qy      615 -----IPLSYSFHLCPQEFPPAAMSYYLDEKGR-----GKHSNCI----- 650
Db      503 ELPGSGDQGSYEFPHLTLQAFPTAFVLVDDRVTQELLRFQEMWMPAGAATSCYPPF 562
Qy      651 -----IDLEKTLBAYGIGHLFGASTTRFLGLSDGE-----R 684
Db      563 LPFOCLQSGPARBDFKXKD---HFQF---TNLFLGCLSKAKQKLRHLVPAALRR 615
Qy      685 EMENTFHCRLSGGRNLMQWPSL-----QLLQPHSLBSLHCLYETRNKFTLQVMAHF 738
Db      616 KKKALMAHLPSSLRGYLKSLEPRVQESFNQVQAMPFTIMLRCTYEQ-----SQKVGOL 670
Qy      739 EEMGMCVETDMLLVCTFC-----IKFSRH--VKKLOL-----IEGRQHSWTSP 781
Db      671 AARGICANY-LKLTVCNACSDCASLSFVHLHFPKRLALDNNNNDVGAHELQCFGR 729
Qy      782 TMYVLFPMVPTDAYWQILFSLVLTNRNKELDLGSNSLSHSAVSKLCTLRPRCLLET 841
Db      730 LTVLRLSVNQITDGGVKVISEELTKYKIVTYGLVNNQITDVGARYVTILDECKGLTH- 788
Qy      842 LRLAGCGLTAEDCKOLAFGLRANQTLTELDSFNVLTDGAKHLQORLR-OPSCKLQRLQ 900
Db      789 LKLGKXKITSBEGKTLALAVKSKSISEVGMNGQVGDGAKAPALNHRHS--LTTLS 846
Qy      901 LVSCGLTSDCCDGLASVLSASPSLKELDLQNNLDDVGRLLCEGLR-HPACKLIRLGLD 959
Db      847 LASNGISTEGKSLARALQONTSLIETLWLTQNELNDEVAESLAEMLKVNQTLKHLMLION 906
Qy      960 QTT-----LSDBRM-----OELRALQOK 978
Db      907 QITAKGTALDALQSNLTGITEICLNGNLKPEAKVYEDEK 948

```

```

RESULT 3
US-09-207-359B-8
; Sequence 8, Application US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT

```

ORGANISM: Homo sapiens
US-09-207-359B-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Pred. No. 3,7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;

237 VVGTPPOAHTSLQPHHWPBVSRESICSTWPKNEDPNOKFTQLLLRPHPSODPLV 296
120 VVNTDPVSRVYQQRHHLGRDS-KFVLC-----YAKR-ELLLEIYMDTIMELY 167
297 KRSPDVEENRGLIIRDLFG--PGIDTQEPRIVLQGAAGIKSTLAROUEAMGRG 354
168 -----GFSNESLGSLSLACLDHTTIGLNEQGETIFILGAGVSKMLQRLQSLWATG 222
355 QL-YGDRFOHVFFSCRELA---QSKVSLAEILGKD--GTATPARIQILSR-PERLL 406
223 RLDAQVGF--FFHRCRMFSCFKESDRICLDLLFKHYCYERBPDEEVAFILRPHVAL 280
407 FILDGVDEPQVNLQEPSSSLCH-----WSQPOPADALLGSLGKTIILPEASFLITA 458
281 FTFDGLD-----ELHSDLDLSRVDPSSCPMBEAPHL-VLLANLISGKILKGSKLITA 332
459 RTTALQNLISLBQARWVVLGFSSESSEKREYFYRFTDERQAIRAFRLVSKNELMALCL 518
333 RTGI---EVPRQPLRKVVLLRGFSPSHRAYARMFPERALQDLISQLEAMPULCSICS 389
519 VPWVSWLACTCLMOQMRKE-----KLTLSKTTTTLCLHY-----LAQALQAP 563
390 VPLFCWIIIFRCQFHRAAFEGSPQLPDCMTLTDVFLVTEVHLNRMQPSLVQRNTRSP 449
564 L-----GPQLRD-LCSL---AAGIOWKTLFSPDDIRKGLDGAITSTLKMGILQENP 614
450 VETLHAG---RDTLCISIGQVAHRGMEKSLVFYTOEEVQASGLQ---ERDMQJLFLRALP 502
615 -----IPLSYFHLGCFQEFFAAMSVYLDEKGR-----GKHSNCT--- 650
503 ELGSGGQOSIEFFHLLQAFFTAFLVLDVDRVGTQELRPFQEMMPAGAAITTSCTPPF 562
651 -----IDLEKTLAAYGHLFGASTTRFLLGLISDEGE-----R 684
563 LRFQCLQSGPARBDELFNKND---HFQF---TNLFLGILSKAKQKLLRHLPAAALRR 615
685 EMENIFHCRLSOGNLMQWVPSL-----QLLQPHSLSLHCLYETRNKTFITQVMAHF 738
616 KRKALMAHLFSSLRGYLSLPRVQVESFNQVQAMPTFIMLRCIYETQ---SQKVGOL 670
739 EEMGMCVETDMEILLVCTFC-----IKFSRH--VKKLQI-----IEGRHSTWSP 781
671 AARGICANY-IKLTTCNACGADCSALSFVLHHPKRLALDDNNNNDYGVRELOPCFSR 729
782 TMVVLFRWVPVTDAYVQIILFSLVKVTRNLKELDLSGNSLSHSAVKSJCKTLRRRCILET 841
730 LTVLRSLVNOITDQGVVLSSEELTKYKIYVYLGLYNNQITDVGARYVTKILDECKGLTH- 788
842 LRLAGCGILADBCDILAFGLRANQTLTFLDLSFVNLTDGAKKHICQIR-OPSCKLOQLQ 900
789 LKLGKKNKITSBGGKYLLAVNKSISIBVGMGQVQVDEGAKFAELRNHPS--LTTLS 846
901 LVSGCLSDCCODLASYLSASPSLKELDLQONNLDVGRVLLCGLR-HRACKILRLGLD 959
847 LANSIGISTEGGKSILARALQONTSLIILWLTONEINDEVAESLAMLKNOQLKRLMLION 906
960 QTT-----LSDEMR-----QELRALDEK 978
907 QITAKGTAQLADALQNSMTGITEICANGNLIKPEAKVYEDEX 948

RESULT 4
US-09-340-620A-8
; Sequence 8, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:

APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/340,620A
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 953
TYPE: PR1
ORGANISM: Homo sapiens
US-09-340-620A-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Pred. No. 3,7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;

237 VVGTPPOAHTSLQPHHWPBVSRESICSTWPKNEDPNOKFTQLLLRPHPSODPLV 296
120 VVNTDPVSRVYQQRHHLGRDS-KFVLC-----YAKR-ELLLEIYMDTIMELY 167
297 KRSPDVEENRGLIIRDLFG--PGIDTQEPRIVLQGAAGIKSTLAROUEAMGRG 354
168 -----GFSNESLGSLSLACLDHTTIGLNEQGETIFILGAGVSKMLQRLQSLWATG 222
355 QL-YGDRFOHVFFSCRELA---QSKVSLAEILGKD--GTATPARIQILSR-PERLL 406
223 RLDAQVGF--FFHRCRMFSCFKESDRICLDLLFKHYCYERBPDEEVAFILRPHVAL 280
407 FILDGVDEPQVNLQEPSSSLCH-----WSQPOPADALLGSLGKTIILPEASFLITA 458
281 FTFDGLD-----ELHSDLDLSRVDPSSCPMBEAPHL-VLLANLISGKILKGSKLITA 332
459 RTTALQNLISLBQARWVVLGFSSESSEKREYFYRFTDERQAIRAFRLVSKNELMALCL 518
333 RTGI---EVPRQPLRKVVLLRGFSPSHRAYARMFPERALQDLISQLEAMPULCSICS 389
519 VPWVSWLACTCLMOQMRKE-----KLTLSKTTTTLCLHY-----LAQALQAP 563
390 VPLFCWIIIFRCQFHRAAFEGSPQLPDCMTLTDVFLVTEVHLNRMQPSLVQRNTRSP 449
564 L-----GPQLRD-LCSL---AAGIOWKTLFSPDDIRKGLDGAITSTLKMGILQENP 614
450 VETLHAG---RDTLCISIGQVAHRGMEKSLVFYTOEEVQASGLQ---ERDMQJLFLRALP 502
615 -----IPLSYFHLGCFQEFFAAMSVYLDEKGR-----GKHSNCT--- 650
503 ELGSGGQOSIEFFHLLQAFFTAFLVLDVDRVGTQELRPFQEMMPAGAAITTSCTPPF 562
651 -----IDLEKTLAAYGHLFGASTTRFLLGLISDEGE-----R 684
563 LRFQCLQSGPARBDELFNKND---HFQF---TNLFLGILSKAKQKLLRHLPAAALRR 615
685 EMENIFHCRLSOGNLMQWVPSL-----QLLQPHSLSLHCLYETRNKTFITQVMAHF 738
616 KRKALMAHLFSSLRGYLSLPRVQVESFNQVQAMPTFIMLRCIYETQ---SQKVGOL 670
739 EEMGMCVETDMEILLVCTFC-----IKFSRH--VKKLQI-----IEGRHSTWSP 781
671 AARGICANY-IKLTTCNACGADCSALSFVLHHPKRLALDDNNNNDYGVRELOPCFSR 729
782 TMVVLFRWVPVTDAYVQIILFSLVKVTRNLKELDLSGNSLSHSAVKSJCKTLRRRCILET 841
730 LTVLRSLVNOITDQGVVLSSEELTKYKIYVYLGLYNNQITDVGARYVTKILDECKGLTH- 788

QY 842 LRLAGCGLTAEBCDCLAFGLRANQTLTEHLSFNVLTDGAKHLCQRLR-QPSCXQLRLQ 900
DB 789 LKGNKKTISEGGKYLALAVKSKSISEVGMGNQVDEGAKAPALRNHPS--LTTLS 846
QY 901 LVSCGLTSDCCODLASVLSAPSLKEIDLQONNLDVGVALLCEGLR-HPACKLIRGLD 959
DB 847 LANSNGSTSGKSLAALQONTLELIMLQNLNEVASELSAEMKKNQTLKHLWILQN 906
QY 960 QTT-----LSDEMR-----OELRALBOEK 978
DB 907 QITAKGTACIADALQNTGTITEICLNGNLKPEAKYEDKEK 948

RESULT 5
US-08-910-731-8
Sequence 8, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-910-731-8

Query Match 4.8%; Score 363.5; DB 2; Length 456;
Best Local Similarity 25.5%; Pred. No. 2,7e-26;
Matches 144; Conservative 54; Mismatches 201; Indels 165; Gaps 15;

QY 470 LEOARWNEVY-----GPFSSSRKCYFYRYFTDEQAIRAPFLVSNKEIWA 515
DB 10 LSPARKTELEPLIOQVEVVALDCCGLTEBECK-----DISALBA-----NSLLE 55
QY 516 LCLVWVSWLACTCLMQQMKRKEKLTLSKTTTTLCLHYLAQALQAPLPQRLDL--CS 573

DB 56 LCL-----RTNELGDAG-----VHVLVQGLQSPTRCKIKLSIQNS 91
QY 574 LAEGIWQKTLFSPDDIRKHLGDAIISTFLKMGILQEHPIPLSYSTHLCFOEFPAAM 633
DB 92 LNEAG-----GVLPSTLASLPTLR-----LHLS-----116
QY 634 SYLVEDEKRGHNSCIIDLEKTLNAYGHGLFGASTRFFL-GILSDGEHEMIFHC 692
DB 117 -----DNPLGAGRLTCEGLLDPCQHLKQLLEYC 147
QY 693 RLSQGRNLMQVPSLQLLQPHLSL-----HCLYETRNKTF 730
DB 148 RLTA-----SCEPLASVLRATRALKEITYSNNDIGEARVUGQGLABDACCQLETRLEN 203
QY 731 LTVMAHFEEMKCYETDELLVCTFCIKPSRHVKQLQIEGRORSTWSPTMVVLFRRV 790
DB 204 CGLTPRANCDCIGIVASQSLRELALSGNKLGDVGMALCPGLHPSRLRTL-----WI 258
QY 791 PVTDAVWQI-----LFSYLYKTRNLKEIDLSGNSLSHSAVSLCKTLRPPCLLET 841
DB 259 -----WECGITAKCGDLCRYLRKESIKELSLAGNELGDEGARLLCETLLBPGQLES 312
QY 842 LRLAGCGLTAEBCDCLAFGLRANQTLTEHLSFNVLTDGAKHLCQRLRQPSCXQLRLQ 901
DB 313 LWKSGSFTIACCPHPSVLAQNRFLLELQISNNRLLEDAGVRELQGLGQPSVLRVWL 372
QY 902 VSCGLTSDCCODLASVLSAPSLKEIDLQONNLDVGVALLCEGLRHPACKLIRGLDOT 961
DB 373 ADCDVSDSSCSLAATLTLNHSIRELDSNNCLGDAGIQLVESYRQPCLLLEQLVLYDI 432
QY 962 TLDSEMRQELRALBOEKPOLLIFS 985
DB 433 YVSEEMEDRLQALBKQKSLRIVIS 456

RESULT 6
US-08-910-731-4
Sequence 4, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003

INFORMATION FOR SEQ ID NO: 6
 SEQUENCE CHARACTERISTICS:
 LENGTH: 461 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 US-08-910-731-6

Query Match 4.7%; Score 357; DB 2; Length 461;
 Best Local Similarity 37.9%; Pred. No. 1.2e-25;
 Matches 86; Conservative 30; Mismatches 91; Indels 20; Gaps 3;

QY 768 QLEGGHSTSTPTWVLFVFWVPTDVAWQI-----LSVAKVTNKLKELDISGN 818
 DB 246 ELCPGLHPSSRRLTL-----WV-----WECGTTAKCGDLRVLRKSKLSKLISLGN 294
 QY 819 SLHSNVSLSKTLRRPRCLLETLRLAGCGLTAECDKDLAFGLRANQTLFELDFSVLT 878
 DB 295 ELDBEGARLLCELTLEGGCQLESIMWVSCSFTTAAACCHSSVLAQNRFLLEQLISNRLE 354
 QY 879 DAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVVG 938
 DB 355 DAGVRELQGLGQPGSVLRYVLTADCDVSDSCSLAATLHSHSLRELDLSNNCLGDAG 414
 QY 939 VLLCGLRHPACKLRLGLDQTTLSDEMQRLEPALEQEPOLLTS 985
 DB 415 ILQLVSVRQPGCLLEQLVLYDIYWESEMDRLQALEKDPRLRVIS 461

RESULT 9

US-08-910-731-2
 Sequence 2, Application US/08910731
 Patent No. 593240

GENERAL INFORMATION:
 APPLICANT: CHATTERJEE, DEB K.
 APPLICANT: SHANDILYA, HARINI
 TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/910,731
 FILING DATE: (Herewith)
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/795,395
 FILING DATE: 04-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/794,546
 FILING DATE: 03-FEB-1997
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: ESMOND, ROBERT W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0942.3440003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-910-731-2

Query Match 4.7%; Score 354.5; DB 2; Length 456;
 Best Local Similarity 34.3%; Pred. No. 2e-25;
 Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

QY 752 LVCTFKRSRHHKQLQLEGGHSTSTPTWVLFVFWVPTDVAWQILFVAKVTNKLK 811
 DB 127 LICEGLDPPQCHRLQL-----FYCRVLRASCEPLASVLRATRAK 168
 QY 812 ELDSLGNLSHSNVSLSKTLRRPRCLLETLRLAGCGLTAECDKDLAFGLRANQ-TLTEL 870
 DB 169 ELTVSNNDIGEGARVLAGGLDASACQLETLLENCGTLPAKCKDLG-DIVASQASRLRL 227
 QY 871 DLFSVLTLAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQ 930
 DB 228 DLGSNGLDAGIAELCPGLLSPASRLKTLWECDTTASGCRDLCKVLQAKETLKELSLA 287
 QY 931 QNNLDVGVRLICEGLRHPACK-----LIRLGLDQTT 962
 DB 288 GNLDBEGARLLCESLLOPGCQLESIMWVSCSLTAAACQHVSLMTQNHLEQLSSNK 347
 QY 963 LQDEMQRLE-RALDEQEPOLLFSRRKPSVMTPTBGLDTGENSN-TSSLKQRLGSEPA 1020
 DB 348 LGDSGIELQALQSGQGTTLRVLC-----LGDCEVTNSGCSL-----A 386
 QY 1021 ASHVAQANIKLIDVS 1035
 DB 387 SLILANRSLRELDLS 401

RESULT 10

US-08-795-395-2
 Sequence 2, Application US/08795395
 Patent No. 5965399

GENERAL INFORMATION:
 APPLICANT: CHATTERJEE, DEB K.
 APPLICANT: SHANDILYA, HARINI
 TITLE OF INVENTION: Cloning and Expression of Rat Liver and
 TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/795,395
 FILING DATE: 04-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: ESMOND, ROBERT W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0942.3440002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-795-395-2

Query Match 4.7%; Score 354.5; DB 2; Length 456;

Best Local Similarity 34.3%; Pred. No. 2e-25;

Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

DB 752 LVCTFCIKSRHVKUQLIGRHRSTWSPMTVVLFRWVPTDAYWQILFSLKVTENLK 811
 127 LTCGGLDPQCHEKLEQL-----EYCRLTAAACCPPLASVLRATRAK 168
 QY 812 ELDLGNSLSHSAVKSLCTLRRCLETLRLAGCGSLTAEDCKDLAFGRANO-TTEL 870
 DB 169 ELTVSNNDIGEMARVGGGLADACQLETLRLKEMCGLTPANCDC-LGVAQASIREL 227
 QY 871 DLSFNVLTDAAGHLCORLRQPSCKLQRLQVSGCLTSDCCODLASVLSASPSIKELDLQ 930
 DB 228 DLGSLGIDGAGIELCEGLISPSRKLTLWMECDITASGCRDLCKVLQKETEISLA 287
 QY 931 QNNLDVGVALLCEGLRHPACK-----LIRLGIDQTT 962
 DB 288 GNTLGGEGARLTCESLIQPCQLESILWVKSCLTAACCGHVSIMLTQNKLLLEQLSSNK 347
 QY 963 LSLBEMQEL-RALBOEKPOLIFSRKRPVMTPEGLDTGEMNS-TSSLKROLGSERA 1020
 DB 348 LGDSGIQELCOALSQPETTLRLVLC-----LGDEVTNSGCCSL-----A 386
 QY 1021 ASHVAQANLKLDPVS 1035
 DB 387 SLILANRSRLRDLIS 401

RESULT 11

US-09-245-281-43
 Sequence 43; Application US/09245281

Patent No. 6369196
 GENERAL INFORMATION:
 APPLICANT: Bertin, John
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
 FILE REFERENCE: 07334/118001
 CURRENT APPLICATION NUMBER: US/09/245,281
 EARLIER FILING DATE: 1999-02-05
 EARLIER APPLICATION NUMBER: US 09/207,359
 EARLIER FILING DATE: 1998-12-08
 EARLIER APPLICATION NUMBER: US 09/099,041
 EARLIER FILING DATE: 1998-06-17
 EARLIER APPLICATION NUMBER: US 09/019,942
 NUMBER OF SEQ ID NOS: 44
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 43
 LENGTH: 953
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-245-281-43

Query Match 4.6%; Score 348; DB 4; Length 953;

Best Local Similarity 21.6%; Pred. No. 3.4e-24;

Matches 199; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

DB 234 MAAVGTTPQAAHSLQPHHHPWBSVRESLCTWPMNEDFNOKFTQLLLQRP----- 287
 103 WLSEIGRSP-----SGLIRKTYVNTDPSVXYTQQLRHQLGRDSKFM 144
 QY 288 --HPRSDPLVKSMP-----YVEENRGLHLEIDLF--GGDLTQERIVYLOGAA 336
 DB 145 LCAQAQKDLLEETVMTLMGLVGFNNENSGSLGDLCLLDHSTGVINRGETVFWGDA 204
 QY 337 GIGKSTLARGVKAEMRGQLYGDRFGHVFFVFSRELA---QSKVSLAEILGKD---GTA 390

DB 205 GVGKSLMLQRLQSLMASGRILTSTA-KPFFHFRMRMSPCFRESMTLSQDLLEFKHFCYREQ 263
 QY 391 TPAPIRQILSR-BERLLFLIDGVD---PGWVQEPSSSELCLHWSOPQPADALLGSLIGK 446
 DB 264 DPEEVFSFLRFPHTALFTFDGDELHSDPDLGRVPS--CCWBPRAHPL-VILLANLLSG 320
 QY 447 TLIPKPSFLITATTALQNLIPSLBOARVVEVGFSSSKKEFYRYFTBERQAIRFRL 506
 DB 321 RLKKGAGKLLTAATGV---EVPKQLLRKVTLLNGFSPSHLRAVARRMFPERTAQEHLLQ 377
 QY 507 VKSNKEIMALCLPWPVSWLACTCIMO-----QMKRK 537
 DB 378 LDANPILCSLCGIPFLCWIIFRCFGHFQVYVESSSOLPDCATLNDVLELVTEVHLNR 437
 QY 538 EKULTSKTTTTLCLHYLAQALQAPQLRDLCSLAAGIWOXKTLTSPDDLKRGDL 597
 DB 438 QPSLSVQNRMS-----PATTLRAG--WRLTHALGEVAHGTDKSLPVPFGQEEVQASKLQ 490
 QY 598 GALTSTPLKMGILLQENP-----IPLSYFTHLCQSFEPFAMSYLDE----- 640
 DB 491 ---EGDLQGLFRLAPDVGPEQOGQYEFPHLTLOAFFTAFLVADKYSTREILRFRE 546
 QY 641 ---KGRKSNCTIIDLEKTLKAYGIRGLFGAS-----TTRFLGLSLDEG 682
 DB 547 WTSPGEATSSSCHSRF-----FSFOCLGSRSLGFPDFRNKHQFQTNLFVGLAKAR 600
 QY 683 EREMENIFHCRLSQGNLMQW-----VPSL-----QLLOPHSLSLHCLY 723
 DB 601 OKLLRQLVKALIRRRKALMAHLFASLRSLKSLPRVGGGFQVQVAMPTFLMMLRCLY 660
 QY 724 ETRNKTFIYVMAHFEEMKCVETDMLVCTCICIFSRHVKUQLIGQHSSTWSPMT 783
 DB 661 ETQ-----SQKVRRLARGI---SADYKLAFCAKNCADCSALSFVHHFHQL----- 706
 QY 784 VVLFWRWVPTDAYWQILFSLVKVTRMLKELDLSGNSLSHSAVSLCKTLRPRCLLETLR 843
 DB 707 -----ALDDNNNINXYGVDELQPCFSR---LTVIR 734
 QY 844 LAGCGLTABEDCKDLAFGLRANQTLTELDSFNVLTDAAGHLCQRLRQPSCK-LQRLQV 902
 DB 735 LSVNQITDNGVKVLCBELFYKIVTEFLGLYNNQITDIGARYVAQIIDE--CRGKLHLKG 792
 QY 903 SCGLTDDCCODLASVLSASPSLELDLQNNLDVGVRLLEGLR-HPACKLRLIGDQT 961
 DB 793 KNRTBSGGKCVALLAVKNSTSVIVGVWGQIDBGAFAEALKHPS--LTTLSLAFN 850
 QY 962 TLDENRQEL-RALBOEKPOLIFSRKRPVMTPEGLDTGEMNSSTSLKROLGSERA 1020
 DB 851 GISPEGGKSLAQALKONTTLTVIW-----LTNKEINDEBAEGFAEMLRVNQT 897
 QY 1021 ASHV 1024
 DB 898 LRHL 901

RESULT 12

US-09-207-359B-43
 Sequence 43; Application US/09207359B

Patent No. 6469140
 GENERAL INFORMATION:
 APPLICANT: Bertin, John
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 FILE REFERENCE: 07334-112001
 CURRENT APPLICATION NUMBER: US/09/207,359B
 CURRENT FILING DATE: 1998-12-08
 PRIOR APPLICATION NUMBER: US 09/099,041
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: US 09/019,942
 PRIOR FILING DATE: 1998-02-06
 NUMBER OF SEQ ID NOS: 47
 SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 43
 ; LENGTH: 953
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-207-359B-43

Query Match 4.6%; Score 348; DB 4; Length 953;
 Best Local Similarity 21.6%; Pred. No. 3.4e-24;
 Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

QY 234 MAAVGTPTQAAISQPHHNPWBSVRESICSTWPKNEDFNQKFTOLLQRP----- 287
 DB 103 WLSEIFSP-----SOLIRKTIWTDVPSRYTQQLRHQGRDSKFM 144
 QY 288 --HPRSDPLVRSMPD-----YBENRGHLEIRDLF--GPGLDTPRIVILQGA 336
 DB 145 LCAQKEDLLBETMTLMTGLVGNENNGISGLDCLDSTGVLENHGETVAFVGA 204
 QY 337 GIGKSTLARKVKAANGOLYGRFOHVFFSCRELA---QSKVSLAEIGKD---GTA 390
 DB 205 GVGKSMLORLQSLWASGRITSTA-KFFHFRCRMFSCEFESDMSLDQLFKHFCYPRQ 263
 QY 391 TPAPRIQLSR--PERLLFLDGVDE--PGWLOPSSSLCWMQOPADALLGSLGK 446
 DB 264 DEEVFSFLRPFHTALFTDGLDELHSDFLSRVDS--CCWEPANPL-VLLANLISG 320
 QY 447 TLIPKASPLITARTALQNLIPSLBOARWVEVLGSESSEKKEFYRYFTDERQAIAPRL 506
 DB 321 RLKGAQKLLTARTGV--EVPRQLRKQVLLRGFSPSHLRAYARMFPERTAOEHLQ 377
 QY 507 VSKNELMALCLPVWVSWLACTCLMO-----OMKRX 537
 DB 378 LDANPRLCSLGVPLFCWIIIFRCQHFQVFEBSQSOLPDCAVTLTDLVLTVEVHLNR 437
 QY 538 EKULTSKTTTTLCLHYLAQALQAPRLGQRLDLSLAEGIWOKKTLESPPDLRKRGD 597
 DB 438 QPSSILQVQRTNRS---PAETLRAG--KRTHLALGVANRGTDKSLFVFGQEVQASKIQ 490
 QY 598 GAIISTFLMGLIOEHP-----IPLSYSPHLCFOEFPAAMSUYLEDE----- 640
 DB 491 ---EGDLOGLFRLALPDVGPBQSGSYEPFHLLTQAFPAFLVADKQSTRELLRFPRE 546
 QY 641 ---KGRKSNCLIDLEKTLBAVIGHLFGAS-----TTRFLGLISDBG 682
 DB 547 WTPSGATSSSCHSSF-----FSFQCLGGRSLRGPDPFRNKHQFTWLFCVGLLAKAR 600
 QY 683 EREMEINIFHCRLSOGNLMQW-----VPSL-----QLLQPHSLIESILCY 723
 DB 601 OKLLROLVPAKILRRKRLAMHLFASLSRYLSLPRVQSGGFNVHAMPTFLMRLCTY 660
 QY 724 ETRNKTFILQVNAHFEEMGCVEITDELLVCTFCIFSRHVKKQLIIBGRHRSSTSPFM 783
 DB 661 ETO---SQKVGRLARGL---SADYKLAFCNACSDCSALSVLHFNHQL----- 706
 QY 784 VVLFWRVPTDAYWQLLFSVLKVTNMLKELDLSGNSLSHSAVKSLSKTLRPRCLLETJR 843
 DB 707 -----ALDDNNNLNDYGVDELPRCPSR---LTVIR 734
 QY 844 LAGCGLTADCKDLAFGLRANQTLTELDSFNVLTADAGAKHLCOURLRPSCK-LQRLQV 902
 DB 735 LSNVQITDVGKVLCEELTKYKIVTEFGLYNNQITDIGARYAQLIDE--CRGLKHLKIG 792
 QY 903 SCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCBGLR--HPACKILRLDQ 961
 DB 793 KKRITSEGGCVAMAYKNSTSYDVGMWGNQIDBEAKAFALKQHP--LTTLSLATN 850
 QY 962 TLSDENKREL-RALBOEKPOLIFSRKRSVMTPTGLDTGEMSNSTSLKORLSERA 1020
 DB 851 GISPEGKSLAQALKQNTTLTVIWM-----LTKNELNDSABEPAEMLAVNQ 897
 QY 1021 ASHV 1024
 DB 898 LRLH 901

RESULT 13
 US-09-340-620A-43
 ; Sequence 437; Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:

; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
 ; FILE REFERENCE: 0734-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; PRIOR APPLICATION NUMBER: US 09/245,281
 ; PRIOR FILING DATE: 1998-02-05
 ; PRIOR APPLICATION NUMBER: US 09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 43
 ; LENGTH: 953
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-340-620A-43

Query Match 4.6%; Score 348; DB 4; Length 953;
 Best Local Similarity 21.6%; Pred. No. 3.4e-24;
 Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

QY 234 MAAVGTPTQAAISQPHHNPWBSVRESICSTWPKNEDFNQKFTOLLQRP----- 287
 DB 103 WLSEIFSP-----SOLIRKTIWTDVPSRYTQQLRHQGRDSKFM 144
 QY 288 --HPRSDPLVRSMPD-----YBENRGHLEIRDLF--GPGLDTPRIVILQGA 336
 DB 145 LCAQKEDLLBETMTLMTGLVGNENNGISGLDCLDSTGVLENHGETVAFVGA 204
 QY 337 GIGKSTLARKVKAANGOLYGRFOHVFFSCRELA---QSKVSLAEIGKD---GTA 390
 DB 205 GVGKSMLORLQSLWASGRITSTA-KFFHFRCRMFSCEFESDMSLDQLFKHFCYPRQ 263
 QY 391 TPAPRIQLSR--PERLLFLDGVDE--PGWLOPSSSLCWMQOPADALLGSLGK 446
 DB 264 DEEVFSFLRPFHTALFTDGLDELHSDFLSRVDS--CCWEPANPL-VLLANLISG 320
 QY 447 TLIPKASPLITARTALQNLIPSLBOARWVEVLGSESSEKKEFYRYFTDERQAIAPRL 506
 DB 321 RLKGAQKLLTARTGV--EVPRQLRKQVLLRGFSPSHLRAYARMFPERTAOEHLQ 377
 QY 507 VSKNELMALCLPVWVSWLACTCLMO-----OMKRX 537
 DB 378 LDANPRLCSLGVPLFCWIIIFRCQHFQVFEBSQSOLPDCAVTLTDLVLTVEVHLNR 437
 QY 538 EKULTSKTTTTLCLHYLAQALQAPRLGQRLDLSLAEGIWOKKTLESPPDLRKRGD 597
 DB 438 QPSSILQVQRTNRS---PAETLRAG--KRTHLALGVANRGTDKSLFVFGQEVQASKIQ 490
 QY 598 GAIISTFLMGLIOEHP-----IPLSYSPHLCFOEFPAAMSUYLEDE----- 640
 DB 491 ---EGDLOGLFRLALPDVGPBQSGSYEPFHLLTQAFPAFLVADKQSTRELLRFPRE 546
 QY 641 ---KGRKSNCLIDLEKTLBAVIGHLFGAS-----TTRFLGLISDBG 682
 DB 547 WTPSGATSSSCHSSF-----FSFQCLGGRSLRGPDPFRNKHQFTWLFCVGLLAKAR 600
 QY 683 EREMEINIFHCRLSOGNLMQW-----VPSL-----QLLQPHSLIESILCY 723
 DB 601 OKLLROLVPAKILRRKRLAMHLFASLSRYLSLPRVQSGGFNVHAMPTFLMRLCTY 660
 QY 724 ETRNKTFILQVNAHFEEMGCVEITDELLVCTFCIFSRHVKKQLIIBGRHRSSTSPFM 783


```

Db      661 ETQ-----SQVVGSLAARGI-----SADYLKLAFCNACGADCSALSPVLMHFRQL----- 706
Qy      784 VLFERWVPTDVAWQILFSLVKTRNLKELDLSGNSLSHSAVKSICKTLRPRCLLETIR 843
Db      707 -----ALDDNNNLNDVGVOSLOPFPSR-----LTVIR 734
Qy      844 LAGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAKAKHLCRLRPSCK-LQRLQLV 902
Db      735 LSVNQITDYGKVCBELTKYKIVTFGLVNNQITDIGARYVAQIDLE--CRGLKHLKLG 792
Qy      903 SCGLTSDCCODLASVLSASPSLKELDLQNNLNDVGVFLCEGLR-HPACKLILGLDQT 961
Db      793 KNRITSEGGKCVAAVKNSTIVGVWGNQIGEGAKAFALKDHP--LTTLSLAFN 850
Qy      962 TLSDMEMROL-RALQOEKPOLIFSRKRPVMTPTGDLTGEMSNSTSLKROHLSERA 1020
Db      851 GISEGGKSLAQKONTTLIVIM-----LTKNELDESACFAEMLRVNOT 897
Qy      1021 ASHV 1024
Db      898 LRHL 901

```

RESULT 14
US-08-519-547A-6
Sequence 6, Application US/08519547A
Patent No. 5994082

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: Proteins Essential for the Expression of
TITLE OF INVENTION: Vertebrate MHC Class II Genes, DNA Sequences Encoding Same
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSER: FISH & NEAVE
STREET: 1251 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10020-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/519,547A
FILING DATE: 25-AUG-1995

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: EP94113378.7
APPLICATION NUMBER: EP94113378.7
FILING DATE: 26-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: HALEY, JAMES F.
REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: VOS-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1130 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-519-547A-6

Query Match

Best Local Similarity

4.64; Score 345; DB 2; Length 1130;
23.28; Pred. No. 9.3e-24;

```

Matches 256; Conservative 134; Mismatches 430; Indels 282; Gaps 49;
Qy      13 LEFKKE-----ELKEFOLLANKAHSRSSSGETPA-----QPEKTSGEVVAS 55
Db      111 LEGSKOIFKHIGDEVIGESMEMPAVGGKSKQRRPPEELPDLKWKMPREPVTYVGS 170
Qy      56 YLVAOYGEQRAWDLALHTWEMQGLRSLC-----AQOAGAGHSPPSPYSPSPHLGSPSQ 111
Db      171 LTVGPVSD-----CSTPLCLPLPALFNOBPAGQGRLEKTDQIPMPFSSSSLSCLNLP 223
Qy      112 TSTAVLMPWHELPACTGQSSRRV-----LRQPLDTGRRRREISALILY 157
Db      224 EGPIQFVPTISTLPHGMQISEAGTVSSIFIVHGEVPOASQVPPSPG-----FTV 274
Qy      158 QALPSPDHSPPSOESPNAPFTSTAVLGWSPPOPSIAPR-----EOAPGTQMPLEDTSG 213
Db      275 HGLPTSPDR--PGSTSPASATDL-----PSMEPALTSANMTEKHTSPQCC--AAG 325
Qy      214 IYTEIREREREKSEKRPMAAVVGTTPQAHTSLQPHHPWEPSPVESLCTWPMKNE 273
Db      326 -----EVSNK-LPKMPEPV--EQFYSLQDTYGA-EPAGPDGI----- 359
Qy      274 FNQKFTQLLLQRPHPRSQDPLVKR--SWPDYVEN--RGHLFIRLDFEGGLDTQEPRI 329
Db      360 ---LVEVDLVQARLERSSSKSLERELATPDMAERQAGGLEVLLAAKEHRRPRETRV 415
Qy      330 VILQAGAGIKSTLARQVKEAMGRGOLYGDYGFVFFYSRELAQ--SKVYSIAELIGMDG 388
Db      416 IAVIGKAGQGSKTWAGAVSRAAMACRL--QYDVFESVPHCLNRPDADGLDQLSLG 473
Qy      389 TATPAPI-----RQILSRPERLLFLDGVDEBGWVLOPSSSELCLHWGQPOPAD--- 437
Db      474 ---PQLVADEVFSHILKRPDRVLTLIDAFEB-----LEAQDGLFSTGCPAPABCSL 525
Qy      438 -ALLGSLGKTIIPBASFLITARTALQNLIPSLQEA-RNVEYLGFESESRKEFYNYFT 495
Db      526 RGLIAGLFQKKLRGCTLLTARPRG--RLVQISKXADALFELSGFMEQAQAQVMKTFE 583
Qy      496 D-----ERQAIRAFRLVSNKELWALCLVPWVWLACTCLMQMKRKEKLTLSKTTTLIC 551
Db      584 SSGMTERQD-RALTLARDRPLLSHSHSPILCAVQOLSEALLBELGDALPS-TLTGLY 641
Qy      552 LHYLAQALQAPLQPOLRDLCSLAAB-GIWQKCTL-----FSPDLRRHGLDGAIIITFLK 606
Db      642 VGLIGRAALDSPGG-ALAEIAKLAMELGRHQSTLQDQPPSADV-----TWAMA 691
Qy      607 MGLLOEHP---IPLSI-STHLCF-QEPPAANSVLEDE-----KRGKHSNCI 650
Db      692 KGLVQHPRAABELAFPSFLQCFGLALMLALSGBIKDELPOYIALTPRKRPYDNL 751
Qy      651 IDEKTEAAYGIGLPGASTTRFLGLLSDG-----EREM 686
Db      752 EGVPRFLA-----GLTPQPARCLGALLGPSAASVDRKQVLAARYIKLQPGTLRKL 806
Qy      687 ENIFHC-----RLSGENLMQWVP-----SLQLLQPSHLE--- 717
Db      807 LEHLHCHAEBAEAGIMQHVQBELPRLSFLGTRLTPDAHYLGKALBAAGQDFSLDRST 866
Qy      718 -----SLHCLYETANKFLVQWNAHFEMKMCVETDM-----ELVYCTPICXF 760
Db      867 GICPSGGLSVIGLSCV--TRFRALSTVALWESLROHGETKLQAAREKFTLEPPFAKS 924
Qy      761 SRHYKCL-QLIEGRORHSTWSPTMVVYL-----FRWVPYTDAY-WQILFSLVKYTR 808
Db      925 LKQVEDGKLVQVQTRRSSSEDTAGELPANRDLKLEFALGPIVSGPAPFKLVIRILAFS 984
Qy      809 NLKELD---LSGNSLSHSAVKSICKTLRPRCLLETIRLAGCGITAECDKDLAFGLRANQ 865
Db      985 SLQHLIDLALSENKIGEGVQSLSATF--POL-----K 1015
Qy      866 TLTELDLSFNVLTDAKAKHLCRLRPSCKLQRLQVSCGLTSDCCODLASVLSASPSLK 925
Db      1016 SLETLNLSQNNITDLGAYKLAELPSLALSLLRLSLYNNCTCDVGAESLARVLPDMVSLR 1075

```

QY 926 ELDLQNNLDVGVRLCEGR 947
 Db 1076 VMDVQYKFTRAQAQJLAASLR 1097

RESULT 15
 US-09-207-359B-47
 ; Sequence 47; Application US/09207359B
 ; Patent No. 6463140
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-112001
 ; CURRENT APPLICATION NUMBER: US/09/207,359B
 ; CURRENT FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 966
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: majority sequence
 ; US-09-207-359B-47

Query Match 4.3%; Score 326.5; DB 4; Length 966;

Best Local Similarity 21.9%; Pred. No. 4.5e-22;

Matches 187; Conservative 113; Mismatches 325; Indels 227; Gaps 31;

Db 624 HLFASLRGYLSLPRVQVGSFNVQAMPTEFLMRLCIYENQ-----SQKVQQLARGI--- 676
 QY 746 ETDMEILLVCTPCTKRSRHVYKQLIEGRQHRSTWSPTMVLLFRWVPTDVAIWOILFSVLK 805
 Db 677 --SADYIKLAFCAACADCSALSFVLHFFHKQL----- 707
 QY 806 VTRNLKELDLSGNSLSHSAVKSLSCTLRPRCLLETTLRAGCGTAEDECKDLAFGLRANQ 865
 Db 708 -----ALDDNNNLDVGVQELQPCPSR---LTVLRLSVNOITDGVAVLSSEILTKYK 757
 QY 866 TLTELDLSFNVLTDAKXHLQRLRQPSCK-----LQRLQVSCG-----LTSDDCCQ 912
 Db 758 IYTFELGVYNNQITDVGARYVAQILDE--CKGLTHLSLVYNNQITDVGAKLGKNTSEGGK 815
 QY 913 DLASVLSASPSLKELDLQNNLDVGVRLCEGR-HPACKLIRLGLDOTTLSDEMRQEL 971
 Db 816 YVALAVKNSTSIYDVGMGNOVGBEAKAFABALKDHPG--LTTLSIASNGISTEGGKSL 873
 QY 972 RALEQKPOLLI 983
 Db 874 AQLAQNTSLTV 885

Search completed: January 29, 2004, 13:48:52
 Job time : 31.1576 secs

QY 271 NEDFNOKTQLLLQRP-----HPRSODPLYKRWSPD-----YVEENRGHLIEIR 315
 Db 122 NTPDVSRYTQQLRHQGRDSKFLCVAQKEDLLLEIYMDTLMGLVGFNSLSLGLLA 181
 QY 316 DLF--GRGLDQREPRYILOGAAGIGKSTLARQVKEAMGRQL-YGDRFOHVFFSCREL 372
 Db 182 CLLDHSTGVINEQETVFLVGDAGVGKSMLLQRLQSLMASGRLTAGAKF--FFHFRCRM 239
 QY 373 A---QSKVSLAELIGKO---GATAPAPIRQILSR--PERLLFLDGVDEPQWLQEPSSSE 425
 Db 240 SCKESRRLSLQDLLEFHFCTPPEODPEEVAFILRPHVALFTPDGLD-----ELHSD 292
 QY 426 LCLH-----WSQPPADALLSLGKTLPEASFLITARTTALQNLIPSLQARWE 477
 Db 293 LDISRVDPSSCPWEPAPHL-VLANLLSGKLKAGKLTARTGV--EVRQLLRKVL 348
 QY 478 VLGFSBSRSREYRYRTDRQAIRAVLYKSNKEMALCLVPWVSWLACTCLMQ----- 532
 Db 349 LRGFSPSHLRAYARMPERRAADHLISQLDANPNLCSLGAPVLCWIIFFRCQHFQAP 408
 QY 533 -----QMKRKEKLTLSKTTTLLC--LHYLAQALQAQPLGP 566
 Db 409 EGSSSQLPDCAVTLITDVFLVTEVHLNRMQPSLVQNRNTRSPATTLHAGRDTLHA----- 463
 QY 567 QLRDLCSLAEGIWQKTLFSPDDLKRGKLDGAIISTFLMGILOEHP-----IPLSYS 620
 Db 464 ---LGEVHARGTDKSLFVFGQEEVQASGLQ---EGDQLGLRALPDVGPBGDSYE 515
 QY 621 FTHLCQEFPAAMSYLEDEKGRK-----HANCITIDLEKTLIAYG 661
 Db 516 FFHLTLQAFPTAFLVADDKVGTQELRFOEWTPSGAASSSCHSFL-----SPQCLG 570
 QY 662 IHGLFGASTTRFLGLLSDGGEREMENI FHCRL---SQGRNLQWQVPSL----- 707
 Db 571 GSGRAGED-----LFXNDHFQFTLVFCGLAKAKQKILQVPAALRRKRKALMA 623
 QY 708 -----QLLQPHSLSLHCLYETFRNKTFLTVQVNAHFEMGMCV 745

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 ; Search time 249.019 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-2
Sequence: 1 MAGANGRLACTYFLKKEE.....HLIMEWEKSKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7534	100.0	1429	10	US-09-996-617-2
2	7534	100.0	1429	10	US-09-931-071-2
3	7534	100.0	1429	12	US-10-028-374-15
4	7534	100.0	1429	12	US-10-183-770-15
5	7534	100.0	1429	15	US-10-028-392-11
6	7502	99.6	1473	10	US-09-388-221-2
7	7488	99.4	1429	12	US-10-028-374-3
8	7488	99.4	1429	12	US-10-183-770-3
9	7364	97.7	1399	10	US-09-388-221-4
10	7332	97.3	1443	10	US-09-388-221-6
11	6344	84.2	1454	10	US-09-388-221-10
12	6174	81.9	1424	10	US-09-388-221-12
13	4072	54.0	764	12	US-10-407-866-92
14	2171.5	28.8	442	11	US-09-895-298-139
15	1382.5	18.4	1027	12	US-10-407-866-68

16	1380	18.3	1061	15	US-10-066-521-18	Sequence 18, Appl
17	1375	18.3	1035	11	US-09-965-621-24	Sequence 24, Appl
18	1375	18.3	1035	11	US-10-407-866-24	Sequence 24, Appl
19	1364	18.1	1004	12	US-10-108-260A-3161	Sequence 3161, Ap
20	1279	17.0	1034	12	US-10-028-374-18	Sequence 18, Appl
21	1279	17.0	1034	12	US-10-132-967-5	Sequence 5, Appl
22	1279	17.0	1034	12	US-10-183-770-18	Sequence 18, Appl
23	1279	17.0	1034	14	US-10-127-516-5	Sequence 5, Appl
24	1279	17.0	1034	14	US-10-027-629-5	Sequence 5, Appl
25	1154	15.3	896	15	US-10-066-521-22	Sequence 22, Appl
26	1037	13.8	994	11	US-09-965-621-16	Sequence 16, Appl
27	1037	13.8	994	12	US-10-407-866-16	Sequence 16, Appl
28	1037	13.8	994	15	US-10-066-521-24	Sequence 24, Appl
29	1008	13.4	919	15	US-10-094-742-2718	Sequence 2718, Ap
30	985.5	13.1	1162	15	US-10-216-645-2	Sequence 2, Appl
31	973.5	12.9	635	12	US-10-407-866-90	Sequence 90, Appl
32	972.5	12.9	858	9	US-09-848-035-8	Sequence 8, Appl
33	972.5	12.9	858	10	US-09-986-224-8	Sequence 8, Appl
34	969	12.9	1143	15	US-10-216-645-4	Sequence 4, Appl
35	948	12.6	1062	12	US-10-239-663-43	Sequence 43, Appl
36	945.5	12.5	1033	12	US-10-132-967-2	Sequence 2, Appl
37	945.5	12.5	1033	14	US-10-127-516-2	Sequence 2, Appl
38	945.5	12.5	1033	14	US-10-027-629-2	Sequence 2, Appl
39	945.5	12.5	1344	15	US-10-066-521-6	Sequence 6, Appl
40	943.5	12.5	1016	10	US-09-986-224-19	Sequence 19, Appl
41	941.5	12.5	1049	12	US-10-239-663-42	Sequence 42, Appl
42	924.5	12.3	732	12	US-10-407-866-72	Sequence 72, Appl
43	923.5	12.2	952	12	US-10-407-866-70	Sequence 70, Appl
44	913.5	12.1	674	12	US-10-407-866-89	Sequence 89, Appl
45	908.5	12.1	980	9	US-09-848-035-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-996-617-2
Sequence 2, Application US/09996617
Patent No. US20020128198A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-340001
CURRENT APPLICATION NUMBER: US/09/996,617
PRIOR FILING DATE: 2001-11-27
PRIOR APPLICATION NUMBER: 09/331,071
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/428,252
PRIOR FILING DATE: 1999-10-27
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1429
TYPE: PRT
ORGANISM: Homo sapiens
US-09-996-617-2

Query Match 100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGANGRLACTYFLKKEEKFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
1 MAGANGRLACTYFLKKEEKFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
QY YGQRAMDALATWEGMGLRSLCAQOEGAGHSPSPYPSSEPHLSPSPSTAYLMPV 120
1 YGQRAMDALATWEGMGLRSLCAQOEGAGHSPSPYPSSEPHLSPSPSTAYLMPV 120
DB YGQRAMDALATWEGMGLRSLCAQOEGAGHSPSPYPSSEPHLSPSPSTAYLMPV 120
1 YGQRAMDALATWEGMGLRSLCAQOEGAGHSPSPYPSSEPHLSPSPSTAYLMPV 120
QY 121 IHLPAQCTGGSSRRVLRQLPDTSGRMRERISALLYLQALPSSPDHSPQESFNAPTST 180
1 IHLPAQCTGGSSRRVLRQLPDTSGRMRERISALLYLQALPSSPDHSPQESFNAPTST 180

```

Db      121 IHLPACTGQSSRRVLRQLPDTSGRRWREISASLYQALPSSPDHESPQSPNAPTST 180
Qy      181 AVIGSGSPPOPSLAPREQAPGTOWPLDSTGIIYITEIRERERESEKGRPPMAAVGT 240
Db      181 AVIGSGSPPOPSLAPREQAPGTOWPLDSTGIIYITEIRERERESEKGRPPMAAVGT 240
Qy      241 PPOAHTSLOPHHPWESPVSRLCSTWPKNEDFNQFTQLLLQRPHPRODPLVKRSW 300
Db      241 PPOAHTSLOPHHPWESPVSRLCSTWPKNEDFNQFTQLLLQRPHPRODPLVKRSW 300
Qy      301 PDVEENRGLIIRDLFGPGLDTPQEPRIYIIQGAAGIGKSTLAROYKAMRGOLYGR 360
Db      301 PDVEENRGLIIRDLFGPGLDTPQEPRIYIIQGAAGIGKSTLAROYKAMRGOLYGR 360
Qy      361 FOHVFFSCRELAQSKVSLAEIIGKGTATPAPIRQILSRPRLFIIDGVEPGMVLQ 420
Db      361 FOHVFFSCRELAQSKVSLAEIIGKGTATPAPIRQILSRPRLFIIDGVEPGMVLQ 420
Qy      421 EPSESLCLHWSQOPADALIGSLGKTIIPBASFLITARTALONLPSLEQARWVVLG 480
Db      421 EPSESLCLHWSQOPADALIGSLGKTIIPBASFLITARTALONLPSLEQARWVVLG 480
Qy      481 FSESSEKEYFYRFTDERQAIRAFRLVKSXKELMALCLVPMVSMGLACTCMQMKREKL 540
Db      481 FSESSEKEYFYRFTDERQAIRAFRLVKSXKELMALCLVPMVSMGLACTCMQMKREKL 540
Qy      541 TLTSKTTTTLCLHYLAQALQAGPLQRLDCLAAEGIQKKTLPSPDLRKHGDLGAI 600
Db      541 TLTSKTTTTLCLHYLAQALQAGPLQRLDCLAAEGIQKKTLPSPDLRKHGDLGAI 600
Qy      601 ISTFLMGLIOEHPILSYSLFIHLCQEPFAAMSYLEDEKGGKXSNCTIIDKLTLEAY 660
Db      601 ISTFLMGLIOEHPILSYSLFIHLCQEPFAAMSYLEDEKGGKXSNCTIIDKLTLEAY 660
Qy      661 GIHGLFGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWPSIQLLQPHSLESLH 720
Db      661 GIHGLFGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWPSIQLLQPHSLESLH 720
Qy      721 CLYETENKTFILQVMAHPEEMGMCVETDMLVCTFCIKFSRHVKKLQILBGRHSTWS 780
Db      721 CLYETENKTFILQVMAHPEEMGMCVETDMLVCTFCIKFSRHVKKLQILBGRHSTWS 780
Qy      781 PMVAVLFRWVPYMDAYQOILFSTLXKTRNLKELDLSGNSLSHSAVSLCTTLARPCLIE 840
Db      781 PMVAVLFRWVPYMDAYQOILFSTLXKTRNLKELDLSGNSLSHSAVSLCTTLARPCLIE 840
Qy      841 TLRACGGLTAEDCKDLAFGLRANQTLTELDELSPVLTDAAGKHLCORLRQPSCKLQRLQ 900
Db      841 TLRACGGLTAEDCKDLAFGLRANQTLTELDELSPVLTDAAGKHLCORLRQPSCKLQRLQ 900
Qy      901 LVSCGLTSDCCODLASVLSASPSLKEBLDLOQNNLDVGVALLCEGLRHAPACKLIRGLDQ 960
Db      901 LVSCGLTSDCCODLASVLSASPSLKEBLDLOQNNLDVGVALLCEGLRHAPACKLIRGLDQ 960
Qy      961 TTLSDBMROELRLBEKPOLLI.FSRKRPVMTPTGDLONGEMSNSTSLKORLISERA 1020
Db      961 TTLSDBMROELRLBEKPOLLI.FSRKRPVMTPTGDLONGEMSNSTSLKORLISERA 1020
Qy      1021 ASHVAQANLKLIDVSKIPIAELAEBSSEPEVVEVELLQVSPASQODLHKRPGLTDDDF 1080
Db      1021 ASHVAQANLKLIDVSKIPIAELAEBSSEPEVVEVELLQVSPASQODLHKRPGLTDDDF 1080
Qy      1081 GPRGPVATEVVDKELKYRHFPVAGSYRWPNTGLCFVMEAVTVEIEFCVMDQFLGEIN 1140
Db      1081 GPRGPVATEVVDKELKYRHFPVAGSYRWPNTGLCFVMEAVTVEIEFCVMDQFLGEIN 1140
Qy      1141 POMSVMVAGPLIITKABGAVEAVHLPHFVALOGGHVDSL.PQAHKKEGMLLEBPARY 1200
Db      1141 POMSVMVAGPLIITKABGAVEAVHLPHFVALOGGHVDSL.PQAHKKEGMLLEBPARY 1200
Qy      1201 ELHHIYLENPSFSPGLVLLKMHNAALRFPVTSVULLYHRVHEBEVTHLYLIPSDCSIR 1260

```

```

Db      1201 ELHHIYLENPSFSPGLVLLKMHNAALRFPVTSVULLYHRVHEBEVTHLYLIPSDCSIR 1260
Qy      1261 KELELCYRSPGDDQLESEFYVGHLSGIRLOVKKDETLVWEALVKPDMLPATTLIP 1320
Db      1261 KELELCYRSPGDDQLESEFYVGHLSGIRLOVKKDETLVWEALVKPDMLPATTLIP 1320
Qy      1321 ARIAVSPPLDAPQLHFNQYREQLIARTSVBEVLDKLGQVLSQOYERVAENTRPS 1380
Db      1321 ARIAVSPPLDAPQLHFNQYREQLIARTSVBEVLDKLGQVLSQOYERVAENTRPS 1380
Qy      1381 QMRKLFSLSQSMDRCKDGLYOALKETHPHLLMELMEKSKKGLPLSS 1429
Db      1381 QMRKLFSLSQSMDRCKDGLYOALKETHPHLLMELMEKSKKGLPLSS 1429

```

```

RESULT 2
US-09-931-071-2
; Sequence 2, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USBS THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PaateSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-2

```

```

Query Match      100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MAGGAWGRACLYLEFLKKELEKEPOLLLANKASRSSSGETPAOPEKTSMEVASYLVNQ 60
Db      1 MAGGAWGRACLYLEFLKKELEKEPOLLLANKASRSSSGETPAOPEKTSMEVASYLVNQ 60
Qy      61 YBQRAMDIALHTWQMGRLSLCAQAOEGAGHSPSEFVSPSEPHLGSPOPTSTAVLMPW 120
Db      61 YBQRAMDIALHTWQMGRLSLCAQAOEGAGHSPSEFVSPSEPHLGSPOPTSTAVLMPW 120
Qy      121 IHLPACTGQSSRRVLRQLPDTSGRRWREISASLYQALPSSPDHESPQSPNAPTST 180
Db      121 IHLPACTGQSSRRVLRQLPDTSGRRWREISASLYQALPSSPDHESPQSPNAPTST 180
Qy      181 AVIGSGSPPOPSLAPREQAPGTOWPLDSTGIIYITEIRERERESEKGRPPMAAVGT 240
Db      181 AVIGSGSPPOPSLAPREQAPGTOWPLDSTGIIYITEIRERERESEKGRPPMAAVGT 240
Qy      241 PPOAHTSLOPHHPWESPVSRLCSTWPKNEDFNQFTQLLLQRPHPRODPLVKRSW 300
Db      241 PPOAHTSLOPHHPWESPVSRLCSTWPKNEDFNQFTQLLLQRPHPRODPLVKRSW 300
Qy      301 PDVEENRGLIIRDLFGPGLDTPQEPRIYIIQGAAGIGKSTLAROYKAMRGOLYGR 360
Db      301 PDVEENRGLIIRDLFGPGLDTPQEPRIYIIQGAAGIGKSTLAROYKAMRGOLYGR 360
Qy      361 FOHVFFSCRELAQSKVSLAEIIGKGTATPAPIRQILSRPRLFIIDGVEPGMVLQ 420
Db      361 FOHVFFSCRELAQSKVSLAEIIGKGTATPAPIRQILSRPRLFIIDGVEPGMVLQ 420
Qy      421 EPSESLCLHWSQOPADALIGSLGKTIIPBASFLITARTALONLPSLEQARWVVLG 480
Db      421 EPSESLCLHWSQOPADALIGSLGKTIIPBASFLITARTALONLPSLEQARWVVLG 480

```

```

QY 481 FSSSRKEYFYRFTDEROAIIRAFRLVKSNEKELMALCLVPVSWLACTCLMOQKREKL 540
DB 481 FSSSRKEYFYRFTDEROAIIRAFRLVKSNEKELMALCLVPVSWLACTCLMOQKREKL 540
QY 541 TLTSKTTTTLCLHYLAQALQAPLGPOLRDCISAAGIWOQKTLFSPDDLRRKGLDGA 600
DB 541 TLTSKTTTTLCLHYLAQALQAPLGPOLRDCISAAGIWOQKTLFSPDDLRRKGLDGA 600
QY 601 ISTFLKKGIILOEHPILPISYSFIHLCTQEPFAANSYVLEDEKGRKSHNCIIDEKTEAY 660
DB 601 ISTFLKKGIILOEHPILPISYSFIHLCTQEPFAANSYVLEDEKGRKSHNCIIDEKTEAY 660
QY 661 GINGLFGASTTRFLGLISDBGEREMENIFHCRLSQRNLMOWPSPQLLOPHSLES 720
DB 661 GINGLFGASTTRFLGLISDBGEREMENIFHCRLSQRNLMOWPSPQLLOPHSLES 720
QY 721 CLYETRNKFTLVQVMAHFEEMGCVETDMEILVCTFCIKFSRHVKCLQLEGRHSTWS 780
DB 721 CLYETRNKFTLVQVMAHFEEMGCVETDMEILVCTFCIKFSRHVKCLQLEGRHSTWS 780
QY 781 PTWVVLFRWVPVTDAYWQILFSLVAKTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840
DB 781 PTWVVLFRWVPVTDAYWQILFSLVAKTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840
QY 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSENVLTDAKXHLCORLROPSCKLRLQ 900
DB 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSENVLTDAKXHLCORLROPSCKLRLQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLHAPACKLIRLGLDQ 960
DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLHAPACKLIRLGLDQ 960
QY 961 TTLSDEMRQELRLAECKPOLLIFSRKPSVMTPTBELDVGEMNSNSTSLKROLSGERA 1020
DB 961 TTLSDEMRQELRLAECKPOLLIFSRKPSVMTPTBELDVGEMNSNSTSLKROLSGERA 1020
QY 1021 ASHVAQANLKLIDVSKIFPIAELAEBSPEVAVELLCPSPASQGLDHLKPLCTDDPF 1080
DB 1021 ASHVAQANLKLIDVSKIFPIAELAEBSPEVAVELLCPSPASQGLDHLKPLCTDDPF 1080
QY 1081 GPTGPVATEVVDKKNLYRVHPVAGSYRWPNTGLCFVMEBAVTEIEFCVMDQFLGEIN 1140
DB 1081 GPTGPVATEVVDKKNLYRVHPVAGSYRWPNTGLCFVMEBAVTEIEFCVMDQFLGEIN 1140
QY 1141 PHSMMWAGLIDDKAPGAVEAVHLPHFVALOGGHDVTSIFQMAHKEGMLLEKARV 1200
DB 1141 PHSMMWAGLIDDKAPGAVEAVHLPHFVALOGGHDVTSIFQMAHKEGMLLEKARV 1200
QY 1201 ELHHIVLENSFSPLGYLKMIMHNAIRFIPVTSVLLYHRVHPREVTFFHLIIPSDCSIR 1260
DB 1201 ELHHIVLENSFSPLGYLKMIMHNAIRFIPVTSVLLYHRVHPREVTFFHLIIPSDCSIR 1260
QY 1261 KELELCYRSPGEBDQFSEFYVGHLSGIRLQVOKDKDETLVWEALVPRGDLMPATLIP 1320
DB 1261 KELELCYRSPGEBDQFSEFYVGHLSGIRLQVOKDKDETLVWEALVPRGDLMPATLIP 1320
QY 1321 ARIAVPBLAPOLMHHVDQYREOLIRAVTSVEVVDKLGQVLSQOQYERVALENTIPS 1380
DB 1321 ARIAVPBLAPOLMHHVDQYREOLIRAVTSVEVVDKLGQVLSQOQYERVALENTIPS 1380
QY 1381 QMRKLFSLSGSWDKCKDGLYQALKETHPHILMELMEKSGSKGLLPLSS 1429
DB 1381 QMRKLFSLSGSWDKCKDGLYQALKETHPHILMELMEKSGSKGLLPLSS 1429

```

RESULT 3
US-10-028-374-15

; Sequence 15, Application US/10028374

; Publication No. US20030143706A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: A NOVEL HUMAN LECITHIN-RICH REPEAT CONTAINING PROTEIN EXPRESSED

; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRBM1

```

; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-15

```

Query Match 100.0%; Score 7534; DB 12; Length 1429;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAGGAMRLACYLEFLKESEIKEFOLLANKASRSSGCTPAQPEKTSQMEVASYVAQ 60
DB 1 MAGGAMRLACYLEFLKESEIKEFOLLANKASRSSGCTPAQPEKTSQMEVASYVAQ 60
QY 61 YGEQRAMDIALHTWQWGLRSLCAQAGEAGHSPPSPYSPSBPHLGSPQPTSTAVLMP 120
DB 61 YGEQRAMDIALHTWQWGLRSLCAQAGEAGHSPPSPYSPSBPHLGSPQPTSTAVLMP 120
QY 121 IHELPAGCTGSRFRVLRLPDTSGRRMRISLSLYQALPSSPDHESPEQSNATPTST 180
DB 121 IHELPAGCTGSRFRVLRLPDTSGRRMRISLSLYQALPSSPDHESPEQSNATPTST 180
QY 181 AVLSGWSPPQPSLAPRECAPGTOWPLDTSIGIYTEIERERSEKSKRPPAAAVGT 240
DB 181 AVLSGWSPPQPSLAPRECAPGTOWPLDTSIGIYTEIERERSEKSKRPPAAAVGT 240
QY 241 PPOAHTSLOPHHHWPBSVRESICSTWPMKNEFPNOFTOLLILORPHPSODPLVRSW 300
DB 241 PPOAHTSLOPHHHWPBSVRESICSTWPMKNEFPNOFTOLLILORPHPSODPLVRSW 300
QY 301 PDVYENRGLIIRDLFGPGLDTPQPRVYILGAGIGSTLAROYKAMGSGQLYGR 360
DB 301 PDVYENRGLIIRDLFGPGLDTPQPRVYILGAGIGSTLAROYKAMGSGQLYGR 360
QY 361 FOHVFYFSCRELAQSKVSLAEILGKGTATPAPIRQILSRPRLILIDGVDEPGVLO 420
DB 361 FOHVFYFSCRELAQSKVSLAEILGKGTATPAPIRQILSRPRLILIDGVDEPGVLO 420
QY 421 EPSEBELCWHSPQPADALIGSLIGKTIIPASFLITARTALONTLPSLEQARWEVLG 480
DB 421 EPSEBELCWHSPQPADALIGSLIGKTIIPASFLITARTALONTLPSLEQARWEVLG 480
QY 481 FSSSRKEYFYRFTDEROAIIRAFRLVKSNEKELMALCLVPVSWLACTCLMOQKREKL 540
DB 481 FSSSRKEYFYRFTDEROAIIRAFRLVKSNEKELMALCLVPVSWLACTCLMOQKREKL 540
QY 541 TLTSKTTTTLCLHYLAQALQAPLGPOLRDCISAAGIWOQKTLFSPDDLRRKGLDGA 600
DB 541 TLTSKTTTTLCLHYLAQALQAPLGPOLRDCISAAGIWOQKTLFSPDDLRRKGLDGA 600
QY 601 ISTFLKKGIILOEHPILPISYSFIHLCTQEPFAANSYVLEDEKGRKSHNCIIDEKTEAY 660
DB 601 ISTFLKKGIILOEHPILPISYSFIHLCTQEPFAANSYVLEDEKGRKSHNCIIDEKTEAY 660
QY 661 GINGLFGASTTRFLGLISDBGEREMENIFHCRLSQRNLMOWPSPQLLOPHSLES 720
DB 661 GINGLFGASTTRFLGLISDBGEREMENIFHCRLSQRNLMOWPSPQLLOPHSLES 720
QY 721 CLYETRNKFTLVQVMAHFEEMGCVETDMEILVCTFCIKFSRHVKCLQLEGRHSTWS 780
DB 721 CLYETRNKFTLVQVMAHFEEMGCVETDMEILVCTFCIKFSRHVKCLQLEGRHSTWS 780
QY 781 PTWVVLFRWVPVTDAYWQILFSLVAKTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840
DB 781 PTWVVLFRWVPVTDAYWQILFSLVAKTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840

```

QY 841 TRLAGCGLTAEBCDCLAFGLRANOTLTLDLSPNVLTDGAKHLCORLRQPSCKLQRLQ 900
 DB 841 TRLAGCGLTAEBCDCLAFGLRANOTLTLDLSPNVLTDGAKHLCORLRQPSCKLQRLQ 900
 QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 QY 961 TLLSDMRQELRALBQEKQOLLIFSRKRSVMTPTGDLTGEMSNSTSLKQRLGSEBA 1020
 DB 961 TLLSDMRQELRALBQEKQOLLIFSRKRSVMTPTGDLTGEMSNSTSLKQRLGSEBA 1020
 QY 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGDLTHTKLTGDDDFW 1080
 DB 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGDLTHTKLTGDDDFW 1080
 QY 1081 GPTGVATTEVDKXENLYRHFPPVAGSYRMPNTGLCFVWRBAVTEIEFCWMDQFLGEIN 1140
 DB 1081 GPTGVATTEVDKXENLYRHFPPVAGSYRMPNTGLCFVWRBAVTEIEFCWMDQFLGEIN 1140
 QY 1141 PSHSMVAGPLLDITAEBAVAVHLPHVVALQGGHVDTSLFQMAFKKEGMILBEPARV 1200
 DB 1141 PSHSMVAGPLLDITAEBAVAVHLPHVVALQGGHVDTSLFQMAFKKEGMILBEPARV 1200
 QY 1201 ELHHTVLENPSPFGLVLLKMTLHNLRFIPTSVVLVYRHPBEVTPHLVLIIPSDCSIR 1260
 DB 1201 ELHHTVLENPSPFGLVLLKMTLHNLRFIPTSVVLVYRHPBEVTPHLVLIIPSDCSIR 1260
 QY 1261 KELBLCYRSPGEDOLFSEFYVHLSGIRLQVCKDETLWEALVYKPDMLPATTLIP 1320
 DB 1261 KELBLCYRSPGEDOLFSEFYVHLSGIRLQVCKDETLWEALVYKPDMLPATTLIP 1320
 QY 1321 ARIAVSPPLDAPQLHFDVQYREOLIAVTSVYVLDKLGQVLSQOYERVLAEINTPS 1380
 DB 1321 ARIAVSPPLDAPQLHFDVQYREOLIAVTSVYVLDKLGQVLSQOYERVLAEINTPS 1380
 QY 1381 QMRKLFSLSQSMDRCKDGLYOALKETHPHLMELMEKSKKGLPLSS 1429
 DB 1381 QMRKLFSLSQSMDRCKDGLYOALKETHPHLMELMEKSKKGLPLSS 1429

RESULT 4
 US-10-183-770-15
 ; Sequence 15, Application US/10183770
 ; Publication No. US20030180812A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Biisecol-Nyera Squabb Company
 ; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
 ; FILE REFERENCE: D00672A CIP
 ; CURRENT APPLICATION NUMBER: US/10/183,770
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: US 60/257,773
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 10/028,374
 ; PRIOR FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1429
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-183-770-15

Query Match 100.0%; Score 7534; DB 12; Length 1429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NAGGAMGRILACTIEPLKKELEKEFQOLLANKAHSSSSSGETPAOERKTSGMVAVSLVAQ 60
 DB 1 NAGGAMGRILACTIEPLKKELEKEFQOLLANKAHSSSSSGETPAOERKTSGMVAVSLVAQ 60
 QY 61 YGGRAMDLAHTTWGMLRSLCAQAGAGHSPPFPYSPSEPHLGSPPQSTSTAVLMPW 120

DB 61 YGGRAMDLAHTTWGMLRSLCAQAGAGHSPPFPYSPSEPHLGSPPQSTSTAVLMPW 120
 QY 121 IHEIPAGCTGSSERRVIRQLPDTSGRWRREISASILYQALPSSPDHSPSOESPAPATST 180
 DB 121 IHEIPAGCTGSSERRVIRQLPDTSGRWRREISASILYQALPSSPDHSPSOESPAPATST 180
 QY 181 AVLSGWSPPQPSLAPRQEAAGCTQWPLDFTSGIYYTIRERERERKSRKPPMAAVGT 240
 DB 181 AVLSGWSPPQPSLAPRQEAAGCTQWPLDFTSGIYYTIRERERERKSRKPPMAAVGT 240
 QY 241 PPOAHTSLQPHHHPBPVRSLSCTWPMKNEDEPNQKTOJLLORPPRPSODPLVKSWS 300
 DB 241 PPOAHTSLQPHHHPBPVRSLSCTWPMKNEDEPNQKTOJLLORPPRPSODPLVKSWS 300
 QY 301 PDYVENRNGHLIEIRDLFGPGLDTPRPIVILQGAAGIGKSTLAROYKXEWAGQGLYQDR 360
 DB 301 PDYVENRNGHLIEIRDLFGPGLDTPRPIVILQGAAGIGKSTLAROYKXEWAGQGLYQDR 360
 QY 361 FOHVYFSCRELAOSKVLSLAELIGKDTATPAPRIQLSSPERLLFLIDGVEDEGWLQ 420
 DB 361 FOHVYFSCRELAOSKVLSLAELIGKDTATPAPRIQLSSPERLLFLIDGVEDEGWLQ 420
 QY 421 EPSSELCLHMSQOPPADLALSLGKTLTPBASFLITARTALQNLISLEQARWEVLG 480
 DB 421 EPSSELCLHMSQOPPADLALSLGKTLTPBASFLITARTALQNLISLEQARWEVLG 480
 QY 481 FSESSEKKEFYRPTDERQAIRAFRLVSKNELMALCVPMVSWLACTCLMQMRKKEKL 540
 DB 481 FSESSEKKEFYRPTDERQAIRAFRLVSKNELMALCVPMVSWLACTCLMQMRKKEKL 540
 QY 541 TLTSKTTTTLCLHTYAQMLQNPPLGQRLDCLSLAEBINQKTLFSPDDLRKHLDDAI 600
 DB 541 TLTSKTTTTLCLHTYAQMLQNPPLGQRLDCLSLAEBINQKTLFSPDDLRKHLDDAI 600
 QY 601 ISTFKMGITLQENHPILPSYSFIHLCFOFFPAMSVLDEDEGRGHSNCIIDLKTLBAY 660
 DB 601 ISTFKMGITLQENHPILPSYSFIHLCFOFFPAMSVLDEDEGRGHSNCIIDLKTLBAY 660
 QY 661 GIHGLFGASTTRFLILGLLSDGEREMENI PHCRLSQGNLMQWPSLQLLQPHLSLSLH 720
 DB 661 GIHGLFGASTTRFLILGLLSDGEREMENI PHCRLSQGNLMQWPSLQLLQPHLSLSLH 720
 QY 721 CLYETRNKTPILQWAAHPEEMCMVETMELLVCFCKEFSRHVKKQLIEGRORSTWS 780
 DB 721 CLYETRNKTPILQWAAHPEEMCMVETMELLVCFCKEFSRHVKKQLIEGRORSTWS 780
 QY 781 PTMVVLFMWVPTDAVYWOILFSVLKVTNKLKELDLSGNSLSHSAVSKLCKTLARPCLE 840
 DB 781 PTMVVLFMWVPTDAVYWOILFSVLKVTNKLKELDLSGNSLSHSAVSKLCKTLARPCLE 840
 QY 841 TRLAGCGLTAEBCDCLAFGLRANOTLTLDLSPNVLTDGAKHLCORLRQPSCKLQRLQ 900
 DB 841 TRLAGCGLTAEBCDCLAFGLRANOTLTLDLSPNVLTDGAKHLCORLRQPSCKLQRLQ 900
 QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLCEGLRHPACKLIRLGLDQ 960
 QY 961 TLLSDMRQELRALBQEKQOLLIFSRKRSVMTPTGDLTGEMSNSTSLKQRLGSEBA 1020
 DB 961 TLLSDMRQELRALBQEKQOLLIFSRKRSVMTPTGDLTGEMSNSTSLKQRLGSEBA 1020
 QY 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGDLTHTKLTGDDDFW 1080
 DB 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGDLTHTKLTGDDDFW 1080
 QY 1081 GPTGVATTEVDKXENLYRHFPPVAGSYRMPNTGLCFVWRBAVTEIEFCWMDQFLGEIN 1140
 DB 1081 GPTGVATTEVDKXENLYRHFPPVAGSYRMPNTGLCFVWRBAVTEIEFCWMDQFLGEIN 1140
 QY 1141 PSHSMVAGPLLDITAEBAVAVHLPHVVALQGGHVDTSLFQMAFKKEGMILBEPARV 1200

```

Db      1141  PSHSMWVAGPLDIDKAEPAVEAVHLPHFVALQGGHVDTSLFQMAHFKESGMLEKPARV 1200
Qy      1201  ELHHIVLENSFSPBLGVLKMHINHALRFIPVTSVLLYHRVHPEEVTFHLLYIPSDCSIR 1260
Db      1201  ELHHIVLENSFSPBLGVLKMHINHALRFIPVTSVLLYHRVHPEEVTFHLLYIPSDCSIR 1260
Qy      1261  KELELCYRSFGEDOLFSEFVYVGHGSGIRLOVOKDKOETLWBEALVKGDLMPATLLIPP 1320
Db      1261  KELELCYRSFGEDOLFSEFVYVGHGSGIRLOVOKDKOETLWBEALVKGDLMPATLLIPP 1320
Qy      1321  ARIAVPSPLDAPOLLHAFVDOYREQLIARVTSVEVLDKLGQVLSQOYERVLAENTRPS 1380
Db      1321  ARIAVPSPLDAPOLLHAFVDOYREQLIARVTSVEVLDKLGQVLSQOYERVLAENTRPS 1380
Qy      1381  QMRKLFSLSSQWDRCKDGLYQALKETHPHLIMELMEKSGKGLPLSS 1429
Db      1381  QMRKLFSLSSQWDRCKDGLYQALKETHPHLIMELMEKSGKGLPLSS 1429

RESULT 5
US-10-028-392-11
; Sequence 11, Application US/10028392
; Publication No. US20030087340A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0085.np
; CURRENT APPLICATION NUMBER: US/10/028,392
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/259,479
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US 60/260,616
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-392-11

Query Match      100.0%; Score 7534; DB 15; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      421  EPSSSLCLHWSQOPADALLGSLGKTLIPASFLITARTTALONLIPSEQARWVEVLG 480
Db      421  EPSSSLCLHWSQOPADALLGSLGKTLIPASFLITARTTALONLIPSEQARWVEVLG 480
Qy      481  FSESSREKEYRFFTERQAIRAFRLVKSNEKEMALCLVWVSWLACTCMMQMKREKL 540
Db      481  FSESSREKEYRFFTERQAIRAFRLVKSNEKEMALCLVWVSWLACTCMMQMKREKL 540
Qy      541  TLTSTKTTTTLCHYLAQALQAPLQPLQDLCSLAEBGWOKTLLPSPDDLRRHGLDGI 600
Db      541  TLTSTKTTTTLCHYLAQALQAPLQPLQDLCSLAEBGWOKTLLPSPDDLRRHGLDGI 600
Qy      601  ISTPLMGILQIEPPIPLASVFHLCQEPFAANSYVLDEKSGKSNCTIDLEKTEAY 660
Db      601  ISTPLMGILQIEPPIPLASVFHLCQEPFAANSYVLDEKSGKSNCTIDLEKTEAY 660
Qy      661  GIGLFGASTTRFLGLLSDGEREMENIFHCRLSGGRNLMQWVPSQLLLQPHLSLELH 720
Db      661  GIGLFGASTTRFLGLLSDGEREMENIFHCRLSGGRNLMQWVPSQLLLQPHLSLELH 720
Qy      721  CLVETRNKTFPLTQVMAHPEEMGCVETDWEILVCTFCIKFSRHVKQLQLEGRQHRSTWS 780
Db      721  CLVETRNKTFPLTQVMAHPEEMGCVETDWEILVCTFCIKFSRHVKQLQLEGRQHRSTWS 780
Qy      781  PTHVVLFRWVPYTDAYQOILFVLTATRNKKEILDGNSLSHSAVKSGLCTTLRRPCLLE 840
Db      781  PTHVVLFRWVPYTDAYQOILFVLTATRNKKEILDGNSLSHSAVKSGLCTTLRRPCLLE 840
Qy      841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSPNVLTDAAGKHLQRLRQPSCKLQRLQ 900
Db      841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSPNVLTDAAGKHLQRLRQPSCKLQRLQ 900
Qy      901  LVSCGLTSDCCODLASYLSASPSLKEILDLOQNNLDVGVALLCEGLHHPACKLIRLGLDQ 960
Db      901  LVSCGLTSDCCODLASYLSASPSLKEILDLOQNNLDVGVALLCEGLHHPACKLIRLGLDQ 960
Qy      961  TTLSDEMRQRLALDEQEKPOLIFSRKPSVMTPTBGLDTGEMSNSTSLKROLSERA 1020
Db      961  TTLSDEMRQRLALDEQEKPOLIFSRKPSVMTPTBGLDTGEMSNSTSLKROLSERA 1020
Qy      1021  ASHVAQANLKLIDVSKIFPIAEIABESSPEVAVVELLCVSPASQGLHTRKPLGTDDDFW 1080
Db      1021  ASHVAQANLKLIDVSKIFPIAEIABESSPEVAVVELLCVSPASQGLHTRKPLGTDDDFW 1080
Qy      1081  GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMEAAVTVEIFCVMQDFLGEIN 1140
Db      1081  GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMEAAVTVEIFCVMQDFLGEIN 1140
Qy      1141  PSHSMWVAGPLDIDKAEPAVEAVHLPHFVALQGGHVDTSLFQMAHFKESGMLEKPARV 1200
Db      1141  PSHSMWVAGPLDIDKAEPAVEAVHLPHFVALQGGHVDTSLFQMAHFKESGMLEKPARV 1200
Qy      1201  ELHHIVLENSFSPBLGVLKMHINHALRFIPVTSVLLYHRVHPEEVTFHLLYIPSDCSIR 1260
Db      1201  ELHHIVLENSFSPBLGVLKMHINHALRFIPVTSVLLYHRVHPEEVTFHLLYIPSDCSIR 1260
Qy      1261  KELELCYRSFGEDOLFSEFVYVGHGSGIRLOVOKDKOETLWBEALVKGDLMPATLLIPP 1320
Db      1261  KELELCYRSFGEDOLFSEFVYVGHGSGIRLOVOKDKOETLWBEALVKGDLMPATLLIPP 1320
Qy      1321  ARIAVPSPLDAPOLLHAFVDOYREQLIARVTSVEVLDKLGQVLSQOYERVLAENTRPS 1380
Db      1321  ARIAVPSPLDAPOLLHAFVDOYREQLIARVTSVEVLDKLGQVLSQOYERVLAENTRPS 1380
Qy      1381  QMRKLFSLSSQWDRCKDGLYQALKETHPHLIMELMEKSGKGLPLSS 1429
Db      1381  QMRKLFSLSSQWDRCKDGLYQALKETHPHLIMELMEKSGKGLPLSS 1429

RESULT 6
US-09-388-221-2
; Sequence 2, Application US/09388221A

```



```

; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; NUMBER OF SEQ. ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-2

```

```

Query Match 99.4%; Score 7502; DB 10; Length 1473;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

```

```

QY 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGTPPAOPEKTSQMEVASTYVAQ 60
DB 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGTPPAOPEKTSQMEVASTYVAQ 60
QY 61 YGEORAWDLAHTWEQWGLSLCAQAQEGAGHSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 120
DB 61 YGEORAWDLAHTWEQWGLSLCAQAQEGAGHSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 120
QY 121 IHELPACCTGSEBRVRLQPLDTSGRMRREISASLYQALPSSPDHSPSOESPNAPTST 180
DB 121 IHELPACCTGSEBRVRLQPLDTSGRMRREISASLYQALPSSPDHSPSOESPNAPTST 180
QY 181 AVUGSWGSPPOPSLAPREQAPGTPQPLDTSQTYTEIERERERESKRRPMAAVVGT 240
DB 181 AVUGSWGSPPOPSLAPREQAPGTPQPLDTSQTYTEIERERERESKRRPMAAVVGT 240
QY 241 PPOAHTSLQHHHPWSPVESLCSQTPWKNEDPNOFTQLLQRPSPSPSPSPSPSPSPSPSPSP 300
DB 241 PPOAHTSLQHHHPWSPVESLCSQTPWKNEDPNOFTQLLQRPSPSPSPSPSPSPSPSPSPSP 300
QY 301 PDVYENRGLIETRLFGPLDTPBRIVILQAGIGKSTLARQYKAMGRGQLYGR 360
DB 301 PDVYENRGLIETRLFGPLDTPBRIVILQAGIGKSTLARQYKAMGRGQLYGR 360
QY 361 FOHVYFSSCELAOSKVSIAELIGKGTATPAPIRILSRPERILFILDGDEPGVTLQ 420
DB 361 FOHVYFSSCELAOSKVSIAELIGKGTATPAPIRILSRPERILFILDGDEPGVTLQ 420
QY 421 EPSEELCIHWSOPPADALGSLIGKTLPEASFLIARFTALONLI PSLQOARWVVLG 480
DB 421 EPSEELCIHWSOPPADALGSLIGKTLPEASFLIARFTALONLI PSLQOARWVVLG 480
QY 481 FSESSREKYFYRYFTDEROAI RARFLVKSNEKLMALCLVPWVSWLACTCIMOQKREKL 540
DB 481 FSESSREKYFYRYFTDEROAI RARFLVKSNEKLMALCLVPWVSWLACTCIMOQKREKL 540
QY 541 TLTSKTTTTCIHLAALQAOPIGPOLBCLSLAAGINQKTLBPDPDLRKGLDGI 600
DB 541 TLTSKTTTTCIHLAALQAOPIGPOLBCLSLAAGINQKTLBPDPDLRKGLDGI 600
QY 601 ISTFLKGILOEHPILSYSPFIHLCPQEPFAANSYVLEDEKRGKSHNCIIDEKTLK 660
DB 601 ISTFLKGILOEHPILSYSPFIHLCPQEPFAANSYVLEDEKRGKSHNCIIDEKTLK 660
QY 661 GIHGLFGASTTRFLGLLSDGGEREMENIFHCRLSQRNLMQWVPSQLLLQPHSLSLH 720
DB 661 GIHGLFGASTTRFLGLLSDGGEREMENIFHCRLSQRNLMQWVPSQLLLQPHSLSLH 720
QY 721 CLVETARKTFLQWMAFEEMGCVETDMELVTFCLIKSRHKKQLLEGQRHSTWS 780
DB 721 CLVETARKTFLQWMAFEEMGCVETDMELVTFCLIKSRHKKQLLEGQRHSTWS 780
QY 781 PTWVVLFRWVPVTDAYQOILFSVLKVTNKLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
DB 781 PTWVVLFRWVPVTDAYQOILFSVLKVTNKLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840

```

```

DB 781 PTWVVLFRWVPVTDAYQOILFSVLKVTNKLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
QY 841 TRLAACGGLYABDCXOLAFGLRANQTLTELDSLFPNLTADAGAKHLCQRLRPSCKLQRLQ 900
DB 841 TRLAACGGLYABDCXOLAFGLRANQTLTELDSLFPNLTADAGAKHLCQRLRPSCKLQRLQ 900
QY 901 LVSCGGLTSPCCODLASVLSASPSLKELDIQQNNLDVQVRLICEGRHHPACKLIRGLQ 960
DB 901 LVSCGGLTSPCCODLASVLSASPSLKELDIQQNNLDVQVRLICEGRHHPACKLIRGLQ 960
QY 961 TTLSDEMROELRALBOEKPOLIFSRKRPVMTPTGELDTGEMSNSTSLKORLSEBA 1020
DB 961 TTLSDEMROELRALBOEKPOLIFSRKRPVMTPTGELDTGEMSNSTSLKORLSEBA 1020
QY 1021 ASHVAQANLKLDDVSKIFPIAIEAESSEVYVVELLCVPSPASQGDHTKPLGTDDEFW 1080
DB 1021 ASHVAQANLKLDDVSKIFPIAIEAESSEVYVVELLCVPSPASQGDHTKPLGTDDEFW 1080
QY 1081 GPTGPVATVVDKELVYVHPVAGSYRWPMTGLCFVNRBAVTVEIEECWDOFLGEIN 1140
DB 1081 GPTGPVATVVDKELVYVHPVAGSYRWPMTGLCFVNRBAVTVEIEECWDOFLGEIN 1140
QY 1141 PQSHWVAGPLDIDIKAPGAVEAVHLPHFVALQGHVDTSLFQMAHFKKEGMLEKPARV 1200
DB 1141 PQSHWVAGPLDIDIKAPGAVEAVHLPHFVALQGHVDTSLFQMAHFKKEGMLEKPARV 1200
QY 1201 ELHHIVLENPSPLGVLKIMINALRPIVTSVVLVYRVRHEEVTFLYILPSCSIR 1260
DB 1201 ELHHIVLENPSPLGVLKIMINALRPIVTSVVLVYRVRHEEVTFLYILPSCSIR 1260
QY 1261 -----KELECYRSGEDOLF 1276
DB 1261 -----KELECYRSGEDOLF 1276
QY 1276 KAIDLEMKFOVFRHKPEPLFPLVWGCYTVSGSSGMLEILPKLELCYRSGEDOLF 1320
DB 1276 KAIDLEMKFOVFRHKPEPLFPLVWGCYTVSGSSGMLEILPKLELCYRSGEDOLF 1320
QY 1321 SEFYVHLSGIRLQYKDKKDETLVWEALYKPGDLMPATTLIPPARIAVPSPLDAPQLH 1336
DB 1321 SEFYVHLSGIRLQYKDKKDETLVWEALYKPGDLMPATTLIPPARIAVPSPLDAPQLH 1336
QY 1337 FVDQYREQLIARTSVYVVLDKLHGQVLSQEOYERVLAEINTPSQNRKLFSLSQSWDRKC 1396
DB 1337 FVDQYREQLIARTSVYVVLDKLHGQVLSQEOYERVLAEINTPSQNRKLFSLSQSWDRKC 1396
QY 1397 KQGLYQALKEHTPHLIMELMEKSGKGLPLSS 1429
DB 1441 KQGLYQALKEHTPHLIMELMEKSGKGLPLSS 1473

```

```

RESULT 7
US-10-028-374-3
; Sequence 3, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ. ID NOS: 22
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-3
Query Match 99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGTPPAOPEKTSQMEVASTYVAQ 60

```

```

Db      1 MAGAGMRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
Qy      61 YGEORANDLALHTWEOMGLRSLCAQAQEGAGHSBSPFYSSEPHLGSPOPTSTAVLMPW 120
Db      61 YGEORANDLALHTWEOMGLRSLCAQAQEGAGHSBSPFYSSEPHLGSPOPTSTAVLMPW 120
Qy      121 IHELPACTGSSERRVRLQPLDTSGRWRREISASILLYALPSSPDHSPQESFNAPTST 180
Db      121 IHELPACTGSSERRVRLQPLDTSGRWRREISASILLYALPSSPDHSPQESFNAPTST 180
Qy      181 AVLSGMSGPPQPSLAPREQAPGTOWPLDTSGLIYTYEIREREREKSEKGRPPMAAVGT 240
Db      181 AVLSGMSGPPQPSLAPREQAPGTOWPLDTSGLIYTYEIREREREKSEKGRPPMAAVGT 240
Qy      241 PPOAHTSLQPHHHPWBSVRESLCSWPMKNEFPNOKFTOLLILQPHPSODPLVGRSW 300
Db      241 PPOAHTSLQPHHHPWBSVRESLCSWPMKNEFPNOKFTOLLILQPHPSODPLVGRSW 300
Qy      301 PDVYENRNGHLEIRDLFGGLDTPREPRIVILQGAAGIGKSTLARQVKAAMGRQOLYGR 360
Db      301 PDVYENRNGHLEIRDLFGGLDTPREPRIVILQGAAGIGKSTLARQVKAAMGRQOLYGR 360
Qy      361 FOHVFYFSCRELAQSKVYSLEILGKDGATAPAPRIOLISPERLFTLDGVDSPGWTLQ 420
Db      361 FOHVFYFSCRELAQSKVYSLEILGKDGATAPAPRIOLISPERLFTLDGVDSPGWTLQ 420
Qy      421 EPSSSLCLHWSQOPADALGSLGKTLTPASFLITARTALONLIPSLQEAWEVTLG 480
Db      421 EPSSSLCLHWSQOPADALGSLGKTLTPASFLITARTALONLIPSLQEAWEVTLG 480
Qy      481 FSESSRKYFRYRYTDEROARLRYKSNELMALCLVPMVSWLACTCLMOMKREKTL 540
Db      481 FSESSRKYFRYRYTDEROARLRYKSNELMALCLVPMVSWLACTCLMOMKREKTL 540
Qy      541 TLTKTTTTCLHYLAOLQAPOLGPOURDCSLAABSTWCKTLFSPDDLRKGLDGA 600
Db      541 TLTKTTTTCLHYLAOLQAPOLGPOURDCSLAABSTWCKTLFSPDDLRKGLDGA 600
Qy      601 ISTEKNGILQENHPIPLSYFIIHLCFOFFAAMSYLEDEKGRKHSNCIIDEKTEAY 660
Db      601 ISTEKNGILQENHPIPLSYFIIHLCFOFFAAMSYLEDEKGRKHSNCIIDEKTEAY 660
Qy      661 GIGHLFGASTRRFLGLISDEGEREMENIFHCRLSOGNRLMOWVPSLQILLQPHLSLH 720
Db      661 GIGHLFGASTRRFLGLISDEGEREMENIFHCRLSOGNRLMOWVPSLQILLQPHLSLH 720
Qy      721 CLVETRKKTFLTOYMAHHEWGMVETDMELLYCTFCIKFSRHAKYKQOLIEGRHSTWS 780
Db      721 CLVETRKKTFLTOYMAHHEWGMVETDMELLYCTFCIKFSRHAKYKQOLIEGRHSTWS 780
Qy      781 PTWVVLFRWVAVTDAYVQILFSLVKTRNLKELDLGNSLSHSAVKSICTLRPRCLLE 840
Db      781 PTWVVLFRWVAVTDAYVQILFSLVKTRNLKELDLGNSLSHSAVKSICTLRPRCLLE 840
Qy      841 TLRLAGCGLTAECDKDLAFGLRANQTLTELDLSFNVLMDAGAKLQRLRQPSCKLQRLQ 900
Db      841 TLRLAGCGLTAECDKDLAFGLRANQTLTELDLSFNVLMDAGAKLQRLRQPSCKLQRLQ 900
Qy      901 LVSGGLSDCCODLASVLSASPSLKEILDQONNLDVGVRLLCGLRHPACKLIRLGLDQ 960
Db      901 LVSGGLSDCCODLASVLSASPSLKEILDQONNLDVGVRLLCGLRHPACKLIRLGLDQ 960
Qy      961 TTLSDEMRQELRALQEKPOLIFSRKRPVMTPTLEGIDTEMNSTSLRKORLGSERA 1020
Db      961 TTLSDEMRQELRALQEKPOLIFSRKRPVMTPTLEGIDTEMNSTSLRKORLGSERA 1020
Qy      1021 ASHVAQANLKLIDVSKIFPIAIEAESSPEVAVELLCPSPASQGLDHTKPLGTDDEFW 1080
Db      1021 ASHVAQANLKLIDVSKIFPIAIEAESSPEVAVELLCPSPASQGLDHTKPLGTDDEFW 1080
Qy      1081 GPFGPVATTEVVDKKNLYRHHFPVAGSYRMPNTGLCFVREAVTVEIEFCWMDQFLGEIN 1140

```

```

Db      1081 GPFGPVATTEVVDKKNLYRHHFPVAGSYRMPNTGLCFVREAVTVEIEFCWMDQFLGEIN 1140
Qy      1141 FOHSMWVAGPLLDIKAEPAVEAVHLPHFVALQGHVDTSLFQWAFREEGMLLEKPARV 1200
Db      1141 FOHSMWVAGPLLDIKAEPAVEAVHLPHFVALQGHVDTSLFQWAFREEGMLLEKPARV 1200
Qy      1201 ELHHTVLENPSFSLGVLKXINHALRPIPTSVYLLYHRVHPEEVTFHLYLISDCSIR 1260
Db      1201 ELHHTVLENPSFSLGVLKXINHALRPIPTSVYLLYHRVHPEEVTFHLYLISDCSIR 1260
Qy      1261 KELELCYRSPEDDLFSEFVYGHGSGIRLOVKOKOBTLVWELVYMGDLMPATTLIP 1320
Db      1261 KELELCYRSPEDDLFSEFVYGHGSGIRLOVKOKOBTLVWELVYMGDLMPATTLIP 1320
Qy      1321 ARIAVSPPLDAPOLLHAFVDOYREQLIARVTSVEVLVDKLHGOVLSSOQYERVLAEINTPS 1380
Db      1321 ACIAVSPPLDAPOLLHAFVDOYREQLIARVTSVEVLVDKLHGOVLSSOQYERVLAEINTPS 1380
Qy      1381 QMRKLFSLQSWDRKCKDGLYALKETHPHILIMELMEKSKKGLPLSS 1429
Db      1381 QMRKLFSLQSWDRKCKDGLYALKETHPHILIMELMEKSKKGLPLSS 1429

RESULT 8
US-10-183-770-3
; Sequence 3, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE OF INVENTION: FREEDOM/INTEL IN BONE MARROW, HLRBM1
; FILE REFERENCE: D0067A CIP
; CURRENT FILING DATE: 2002-06-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-3

Query Match      99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      1 MAGAGMRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
Db      1 MAGAGMRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
Qy      61 YGEORANDLALHTWEOMGLRSLCAQAQEGAGHSBSPFYSSEPHLGSPOPTSTAVLMPW 120
Db      61 YGEORANDLALHTWEOMGLRSLCAQAQEGAGHSBSPFYSSEPHLGSPOPTSTAVLMPW 120
Qy      121 IHELPACTGSSERRVRLQPLDTSGRWRREISASILLYALPSSPDHSPQESFNAPTST 180
Db      121 IHELPACTGSSERRVRLQPLDTSGRWRREISASILLYALPSSPDHSPQESFNAPTST 180
Qy      181 AVLSGMSGPPQPSLAPREQAPGTOWPLDTSGLIYTYEIREREREKSEKGRPPMAAVGT 240
Db      181 AVLSGMSGPPQPSLAPREQAPGTOWPLDTSGLIYTYEIREREREKSEKGRPPMAAVGT 240
Qy      241 PPOAHTSLQPHHHPWBSVRESLCSWPMKNEFPNOKFTOLLILQPHPSODPLVGRSW 300
Db      241 PPOAHTSLQPHHHPWBSVRESLCSWPMKNEFPNOKFTOLLILQPHPSODPLVGRSW 300
Qy      301 PDVYENRNGHLEIRDLFGGLDTPREPRIVILQGAAGIGKSTLARQVKAAMGRQOLYGR 360
Db      301 PDVYENRNGHLEIRDLFGGLDTPREPRIVILQGAAGIGKSTLARQVKAAMGRQOLYGR 360

```

QY 361 FOHVYFSCRELAOSKVSLAELIGKDGATPAPIRQILSRERLLFIIDGVDGEMVQ 420
DB 361 FOHVYFSCRELAOSKVSLAELIGKDGATPAPIRQILSRERLLFIIDGVDGEMVQ 420
QY 421 BPSSELCIHWSPQOPADALIGSLGKTIIPASFLITANTTALONLPSLEQARWEVIG 480
DB 421 BPSSELCIHWSPQOPADALIGSLGKTIIPASFLITANTTALONLPSLEQARWEVIG 480
QY 481 PESSRKEKEXFYRYFTDEROAIRAFRLVKSNEKLMALCLVPWVSWLACTCLMOOMKREKL 540
DB 481 PESSRKEKEXFYRYFTDEROAIRAFRLVKSNEKLMALCLVPWVSWLACTCLMOOMKREKL 540
QY 541 TLTSTKTTTLCIHYLAQALQAPLGPOLRDLCSLAEBGWOKKTLFSPDRLRKHGIDAI 600
DB 541 TLTSTKTTTLCIHYLAQALQAPLGPOLRDLCSLAEBGWOKKTLFSPDRLRKHGIDAI 600
QY 601 ISTFLKMGILQEHPIPLSYSTFHLCOEFPFAMSVYLBDEKGRGHSNCIIDLKTLLEY 660
DB 601 ISTFLKMGILQEHPIPLSYSTFHLCOEFPFAMSVYLBDEKGRGHSNCIIDLKTLLEY 660
QY 661 GIHGFAGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWVPSLOLLQPHLESIAH 720
DB 661 GIHGFAGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWVPSLOLLQPHLESIAH 720
QY 721 CLYETRNKTLFLOVNAHFEEMGCVETDMELLVCTFCIKFSRHVKKLOLIEGRORHSTWS 780
DB 721 CLYETRNKTLFLOVNAHFEEMGCVETDMELLVCTFCIKFSRHVKKLOLIEGRORHSTWS 780
QY 781 PFMVVLFRVVPVTDAYWOILFSLVKTRNLKELIDSGNSLSHSAVSKLCTLRPRCLLE 840
DB 781 PFMVVLFRVVPVTDAYWOILFSLVKTRNLKELIDSGNSLSHSAVSKLCTLRPRCLLE 840
QY 841 TRLAGCGIATBDCDOLAFGLRANOTLTLELDSFNVLTDAGAKHLQORLROPSCKLQRLQ 900
DB 841 TRLAGCGIATBDCDOLAFGLRANOTLTLELDSFNVLTDAGAKHLQORLROPSCKLQRLQ 900
QY 901 IYVSCGLTSDCCDOLASVLSASPSLKELDLQONNLDVGRLLCEGLRHPACKLIRLGLDQ 960
DB 901 IYVSCGLTSDCCDOLASVLSASPSLKELDLQONNLDVGRLLCEGLRHPACKLIRLGLDQ 960
QY 961 TLTSEMREBELALBEKPOLLIISRKRKSVMTPTBGLDTGEMSNSTSLKORLSEBA 1020
DB 961 TLTSEMREBELALBEKPOLLIISRKRKSVMTPTBGLDTGEMSNSTSLKORLSEBA 1020
QY 1021 ASHVANOANLKLIDVSKIFPIAIAIESSEPVVVELLCVPSASQODLTKPLGTDDDPW 1080
DB 1021 ASHVANOANLKLIDVSKIFPIAIAIESSEPVVVELLCVPSASQODLTKPLGTDDDPW 1080
QY 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRMPNTGLCFVAREAVTVEIEFCVMDQFLGEIN 1140
DB 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRMPNTGLCFVAREAVTVEIEFCVMDQFLGEIN 1140
QY 1141 PSHMMVAVPRLDIDYKREBAVEAVHLPHFVALQGGVNDTSLFQVAFKKEGMLLEKPAV 1200
DB 1141 PSHMMVAVPRLDIDYKREBAVEAVHLPHFVALQGGVNDTSLFQVAFKKEGMLLEKPAV 1200
QY 1201 ELHHIYLENPSFSPGLVLLKMHINLRFIPVTSVULLYHRLHPEEVTFLYILPSDCSIR 1260
DB 1201 ELHHIYLENPSFSPGLVLLKMHINLRFIPVTSVULLYHRLHPEEVTFLYILPSDCSIR 1260
QY 1261 KELELCYRSPGEBDQFSEFYVGHLSGIRLOVKKDETLVWEALVKPDLMPATTLIP 1320
DB 1261 KELELCYRSPGEBDQFSEFYVGHLSGIRLOVKKDETLVWEALVKPDLMPATTLIP 1320
QY 1321 ARIAVSPRLDAPOLAHFVDOYREQLIARVTSVEVVDKXGCVLSEOEYERVAENTRS 1380
DB 1321 ARIAVSPRLDAPOLAHFVDOYREQLIARVTSVEVVDKXGCVLSEOEYERVAENTRS 1380
QY 1381 QNRKTLFSLQSWDRCKDGLYQALKETHPHLIMELMEKSKKGLPLSS 1429
DB 1381 QNRKTLFSLQSWDRCKDGLYQALKETHPHLIMELMEKSKKGLPLSS 1429

RESULT 9
US-09-388-221-4
; Sequence 4, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1 Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-4

Query Match 97.7%; Score 7364; DB 10; Length 1399;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 30; Gaps 1;

QY 1 MAGGAWGRACYLEFLKKEBELKEPOLLIANKASHSSSGETPAOEBKTSQMEVASTVLAQ 60
DB 1 MAGGAWGRACYLEFLKKEBELKEPOLLIANKASHSSSGETPAOEBKTSQMEVASTVLAQ 60
QY 61 YGEORAMDLAHTWQMGRLSLCAQAOBAGHSPSPSPSPHIGSPSPQPTSTAVLMPW 120
DB 61 YGEORAMDLAHTWQMGRLSLCAQAOBAGHSPSPSPSPHIGSPSPQPTSTAVLMPW 120
QY 121 IHELPAGCTGSESERVLRLPDTSGRRWRREISASLLYQALPSSPDHESPQSPNAPTST 180
DB 121 IHELPAGCTGSESERVLRLPDTSGRRWRREISASLLYQALPSSPDHESPQSPNAPTST 180
QY 181 AVLSGWSGPPQPSLAPROQAPGTQWPLDETSGIYTYEIRERERKSKSEKGRPMNAVGT 240
DB 181 AVLSGWSGPPQPSLAPROQAPGTQWPLDETSGIYTYEIRERERKSKSEKGRPMNAVGT 240
QY 241 PPOATSLQPHHHPESPVRSLSGTWPWKNDFFNOKFTOLLQRPHPRSODPLVKRSM 300
DB 241 PPOATSLQPHHHPESPVRSLSGTWPWKNDFFNOKFTOLLQRPHPRSODPLVKRSM 300
QY 301 PDYVENRNGHLIEIRDLFGPGLDTPRIVILQGAAGIGKSTLARBQVXAMGRGQLYGBR 360
DB 301 PDYVENRNGHLIEIRDLFGPGLDTPRIVILQGAAGIGKSTLARBQVXAMGRGQLYGBR 360
QY 361 FOHVYFSCRELAOSKVSLAELIGKDGATPAPIRQILSRERLLFIIDGVDGEMVQ 420
DB 361 FOHVYFSCRELAOSKVSLAELIGKDGATPAPIRQILSRERLLFIIDGVDGEMVQ 420
QY 421 BPSSELCIHWSPQOPADALIGSLGKTIIPASFLITANTTALONLPSLEQARWEVIG 480
DB 421 BPSSELCIHWSPQOPADALIGSLGKTIIPASFLITANTTALONLPSLEQARWEVIG 480
QY 481 PESSRKEKEXFYRYFTDEROAIRAFRLVKSNEKLMALCLVPWVSWLACTCLMOOMKREKL 540
DB 481 PESSRKEKEXFYRYFTDEROAIRAFRLVKSNEKLMALCLVPWVSWLACTCLMOOMKREKL 540
QY 541 TLTSTKTTTLCIHYLAQALQAPLGPOLRDLCSLAEBGWOKKTLFSPDRLRKHGIDAI 600
DB 541 TLTSTKTTTLCIHYLAQALQAPLGPOLRDLCSLAEBGWOKKTLFSPDRLRKHGIDAI 600
QY 601 ISTFLKMGILQEHPIPLSYSTFHLCOEFPFAMSVYLBDEKGRGHSNCIIDLKTLLEY 660
DB 601 ISTFLKMGILQEHPIPLSYSTFHLCOEFPFAMSVYLBDEKGRGHSNCIIDLKTLLEY 660
QY 661 GIHGFAGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWVPSLOLLQPHLESIAH 720
DB 661 GIHGFAGASTTRFLGLISDEGEREMENIFHCRLSQGRNLMQWVPSLOLLQPHLESIAH 720
QY 721 CLYETRNKTLFLOVNAHFEEMGCVETDMELLVCTFCIKFSRHVKKLOLIEGRORHSTWS 780
DB 721 CLYETRNKTLFLOVNAHFEEMGCVETDMELLVCTFCIKFSRHVKKLOLIEGRORHSTWS 780

```

Db 721 CLYETRNKFTLQVMAHFEEMGCVETDMELLVCTFCIKFSRHVKQLIIEGRQHRSTWS 780
QY 781 PTMVVLFRRWVPTDAYWQILFVYLKVTNRNKELDLGSNSLSHSAVKSCTTLRRPRCLLE 840
Db 781 PTMVVLFRRWVPTDAYWQILFVYLKVTNRNKELDLGSNSLSHSAVKSCTTLRRPRCLLE 840
QY 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSEFNVLTDAAGAKHLCORLQPSCKLQRLQ 900
Db 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSEFNVLTDAAGAKHLCORLQPSCKLQRLQ 900
QY 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
QY 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
QY 961 TTLSDEMRQELRALBOEKPOLLIFSRKKPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 1020
Db 961 TTLSDEMRQELRALBOEKPOLLIFSRKKPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 1020
QY 958 -----KPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 990
Db 1021 ASHVAQANLKLIDVSKIFPIAETAEBSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1080
QY 991 ASHVAQANLKLIDVSKIFPIAETAEBSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1050
Db 1081 GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMBEAVTVEIFCVMDOFLGEIN 1140
QY 1051 GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMBEAVTVEIFCVMDOFLGEIN 1110
Db 1141 PHSMMVWAGPLDIDIKAPGAVEAVHLPHFVALOGGHVDTSLFQMAHFKESGMLLEKPARV 1200
QY 1111 PHSMMVWAGPLDIDIKAPGAVEAVHLPHFVALOGGHVDTSLFQMAHFKESGMLLEKPARV 1170
Db 1201 ELHHIVLENSFSFPLGVLKMHINALRFIPVTSVLLYHNVHPEVPHLYLIPSDCSIR 1260
QY 1171 ELHHIVLENSFSFPLGVLKMHINALRFIPVTSVLLYHNVHPEVPHLYLIPSDCSIR 1230
Db 1261 KELELCYRSGEDQLFSEFVYVGHLSGIRLQVYDKDETLVWEALVYKPDLMPATTLIIP 1320
QY 1231 KELELCYRSGEDQLFSEFVYVGHLSGIRLQVYDKDETLVWEALVYKPDLMPATTLIIP 1290
Db 1321 ARIAVBPPLDAPOLAHFVDDYREQLIARVTSVEVLDKLGQVLSQOQYERVALENTRPS 1380
QY 1291 ARIAVBPPLDAPOLAHFVDDYREQLIARVTSVEVLDKLGQVLSQOQYERVALENTRPS 1350
Db 1381 QMRKLFSLSSWDRCKDGLYQALKETHPHLIMELMEKSGSKGLLPLSS 1429
QY 1351 QMRKLFSLSSWDRCKDGLYQALKETHPHLIMELMEKSGSKGLLPLSS 1399

```

RESULT 10

```

US-09-388-221-6
; Sequence 6, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1 Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-Lf 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-6

```

```

Query Match 97.3%; Score 7332; DB 10; Length 1443;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 74; Gaps 2;

```

```

QY 1 MAGGAMRCLACYLEFLKKEBELKEFOLLANKAHSRSSGGTTPAPQETKSGMEVASIYVAQ 60
Db 1 MAGGAMRCLACYLEFLKKEBELKEFOLLANKAHSRSSGGTTPAPQETKSGMEVASIYVAQ 60

```

```

QY 61 YGEORAMDALATHTEOMGLRSLCAOAOEGAGHSPFPYSSSEPHLGSPQPTSTAVLMPW 120
Db 61 YGEORAMDALATHTEOMGLRSLCAOAOEGAGHSPFPYSSSEPHLGSPQPTSTAVLMPW 120
QY 121 IHELPACTOGSERVYLROLPTDSGRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
Db 121 IHELPACTOGSERVYLROLPTDSGRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
QY 181 AVIGSWGSPPOPSLARREQOAPCTOWPLDGTSGITYTETREERRESEKRRPPMAAVGT 240
Db 181 AVIGSWGSPPOPSLARREQOAPCTOWPLDGTSGITYTETREERRESEKRRPPMAAVGT 240
QY 241 PPOAHTSLOPHHMPWESVRESLCTWPMKNEPNOKFTOLLILORHPASOPPLVRSW 300
Db 241 PPOAHTSLOPHHMPWESVRESLCTWPMKNEPNOKFTOLLILORHPASOPPLVRSW 300
QY 301 PDVYENRGLHIBIRDLFGPDLTQBPRIYILOGAAGIGKSTLARQYKAMGRGOLYGR 360
Db 301 PDVYENRGLHIBIRDLFGPDLTQBPRIYILOGAAGIGKSTLARQYKAMGRGOLYGR 360
QY 361 FOHVFFPSCHELAOSKVSLAELIGKDGRTTPAPRIQILSRPERLLFILDGVBPGVQLQ 420
Db 361 FOHVFFPSCHELAOSKVSLAELIGKDGRTTPAPRIQILSRPERLLFILDGVBPGVQLQ 420
QY 421 BPSSSELCLHWSQOPADALGSLIGKTIILPEASFLTARTTALONLPSLEQARWVVLG 480
Db 421 BPSSSELCLHWSQOPADALGSLIGKTIILPEASFLTARTTALONLPSLEQARWVVLG 480
QY 481 FSSSSREKEYRYFTDERQAIRAFLVKSNEKELMALCLVWVSWLACTCLMOQKREKL 540
Db 481 FSSSSREKEYRYFTDERQAIRAFLVKSNEKELMALCLVWVSWLACTCLMOQKREKL 540
QY 541 TUSKTKTTTCLHYLAQALQAOPLGPORLDGLAABGIVQKTLTPSPDRLRKHGLDGI 600
Db 541 TUSKTKTTTCLHYLAQALQAOPLGPORLDGLAABGIVQKTLTPSPDRLRKHGLDGI 600
QY 541 TUSKTKTTTCLHYLAQALQAOPLGPORLDGLAABGIVQKTLTPSPDRLRKHGLDGI 600
Db 541 TUSKTKTTTCLHYLAQALQAOPLGPORLDGLAABGIVQKTLTPSPDRLRKHGLDGI 600
QY 601 ISTFELKMGILQOEPIPLSYSFHLCFOEFPAANSVYLEDKRGKSNCTIIDLEKTEAY 660
Db 601 ISTFELKMGILQOEPIPLSYSFHLCFOEFPAANSVYLEDKRGKSNCTIIDLEKTEAY 660
QY 661 GIGHLFGASTTRFLLGLSDGEREMENIFHCHLSQGRNLMQWVPSLQILLPHSLESLH 720
Db 661 GIGHLFGASTTRFLLGLSDGEREMENIFHCHLSQGRNLMQWVPSLQILLPHSLESLH 720
QY 721 CLYETRNKFTLQVMAHFEEMGCVETDMELLVCTFCIKFSRHVKQLIIEGRQHRSTWS 780
Db 721 CLYETRNKFTLQVMAHFEEMGCVETDMELLVCTFCIKFSRHVKQLIIEGRQHRSTWS 780
QY 781 PTMVVLFRRWVPTDAYWQILFVYLKVTNRNKELDLGSNSLSHSAVKSCTTLRRPRCLLE 840
Db 781 PTMVVLFRRWVPTDAYWQILFVYLKVTNRNKELDLGSNSLSHSAVKSCTTLRRPRCLLE 840
QY 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSEFNVLTDAAGAKHLCORLQPSCKLQRLQ 900
Db 841 TLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSEFNVLTDAAGAKHLCORLQPSCKLQRLQ 900
QY 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
QY 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASYLASPSLKELDLQONNLDVGVRLLCEGLRHAPACKLIRGLDQ 960
QY 961 TTLSDEMRQELRALBOEKPOLLIFSRKKPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 1020
Db 961 TTLSDEMRQELRALBOEKPOLLIFSRKKPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 1020
QY 958 -----KPSVMTPTBGLDTGEMSNSTSLKRORLGSERA 990
Db 1021 ASHVAQANLKLIDVSKIFPIAETAEBSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1080
QY 991 ASHVAQANLKLIDVSKIFPIAETAEBSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1050
Db 1081 GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMBEAVTVEIFCVMDOFLGEIN 1140
QY 1051 GPTGPVATEVVDKKNLYRVHFPVAGSYRMPNTGLCFVMBEAVTVEIFCVMDOFLGEIN 1110
Db 1141 PHSMMVWAGPLDIDIKAPGAVEAVHLPHFVALOGGHVDTSLFQMAHFKESGMLLEKPARV 1200

```

```

DB 1111 PSHWVWAPBLDDIKAPAVEVHLPHFVALOGGVDTSLFQMAHFKKEGMLLEKPAIV 1170
QY 1201 ELHHIYLEKPSFSLGVLKMTINLRFIPVTSVLLYHVRHEEYTPHLYLIPSDCSIR 1260
DB 1172 ELHHIYLEKPSFSLGVLKMTINLRFIPVTSVLLYHVRHEEYTPHLYLIPSDCSIR 1230
QY 1261 -----KELCYRSPGEDOLF 1276
DB 1231 KALIDLEMKFQVRIHKPPPLFPLVWGCRTYVSGSGSMLEILPKLELCYRSPGEDOLF 1290
QY 1277 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDMLPATLLIPPAIIVPSPLDAPQLH 1336
DB 1291 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDMLPATLLIPPAIIVPSPLDAPQLH 1350
QY 1337 FVDQYREQLIARYTSVEVLDKLGVLSSOBYERYLAENTRPSQMKLFSLISQSWDRK 1396
DB 1351 FVDQYREQLIARYTSVEVLDKLGVLSSOBYERYLAENTRPSQMKLFSLISQSWDRK 1410
QY 1397 KDLGYOALKETHPHLIMELMEKSKKGLPLSS 1429
DB 1411 KDLGYOALKETHPHLIMELMEKSKKGLPLSS 1443

```

RESULT 11

```

US-09-388-221-10
; Sequence 10, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; FILE OF INVENTION: No. US20020192643A1 Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1454
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-388-221-10

```

Query Match 84.2%; Score 6344; DB 10; Length 1454;
 Best Local Similarity 84.2%; Pred. No. 0;
 Matches 1229; Conservative 53; Mismatches 122; Indels 56; Gaps 7;

```

QY 1 MAGGANGRLACTYELPKELKEFOLLANKAHSSSGGTPAOPKTSQMEVASTLVNQ 60
DB 1 MAGGANGRLACTYELPKELKEFOLLANKAHSSSGGTPAOPKTSQMEVASTLVNQ 60
QY 61 YGGRAMDLALHTWBMGLRSLCAOAGAGHSPPSPSPBPHGSPGSPSTAVLMPW 120
DB 61 YGGRAMDLALHTWBMGLRSLCAOAGAGHSPPSPSPBPHGSPGSPSTAVLMPW 120
QY 121 IHELPACTQGSERRVLRQLPDTSGRRMEISASLLYOLPSSPDHESPQSPNAPTST 180
DB 121 IHELPACTQGSERRVLRQLPDTSGRRMEISASLLYOLPSSPDHESPQSPNAPTST 180
QY 121 IHELPACTQGSERRVLRQLPDTSGRRMEISASLLYOLPSSPDHESPQSPNAPTST 180
DB 121 IHELPACTQGSERRVLRQLPDTSGRRMEISASLLYOLPSSPDHESPQSPNAPTST 180
QY 181 AVLSGSGSPPOPSLARPOAPGTOWPLDETSGIYTYTERERERESKGRPPMAAVGT 240
DB 181 AVLSGSGSPPOPSLARPOAPGTOWPLDETSGIYTYTERERERESKGRPPMAAVGT 240
QY 241 PROAHSLQPHHHPSPVRESLCTWPMKNEDFNKFTOLLQRPHPSPQDPLVKRSW 300
DB 241 PROAHSLQPHHHPSPVRESLCTWPMKNEDFNKFTOLLQRPHPSPQDPLVKRSW 300
QY 301 PDVEENRGLIIRLDLFGGLDTQSPRIYILQGAAGIGKSTLAROYKEMAGRGOLYGR 360
DB 301 PDVEENRGLIIRLDLFGGLDTQSPRIYILQGAAGIGKSTLAROYKEMAGRGOLYGR 360
QY 361 FQHVFFSCRELAQSKVSLAEILGDKGTATPAPRIQLSRPRLFLIDGVDPGWLD 420

```

```

DB 361 FQHVFFSCRELAQSKVSLAEILGDKGTATPAPRIQLSRPRLFLIDGVDPGWLD 420
QY 421 BPSSELCHMSQPOPADLIGSLGKTLIPASFLITARTTLOMLIPSEARWAVYLG 480
DB 421 BPSSELCHMSQPOPADLIGSLGKTLIPASFLITARTTLOMLIPSEARWAVYLG 480
QY 481 PESSRKEFYRYFDEROQAIRAFVLVSNKELMALCVPNVSWLACTLMOQMKREKL 540
DB 481 PESSRKEFYRYFDEROQAIRAFVLVSNKELMALCVPNVSWLACTLMOQMKREKL 540
QY 541 TITSKTTTTLCHAYLAQALQAPLPOLBDCSLAEGIWOKTTFSPDDLARKGLDGA 600
DB 541 TITSKTTTTLCHAYLAQALQAPLPOLBDCSLAEGIWOKTTFSPDDLARKGLDGA 600
QY 601 ISTFLKMGILQEHPLPLSYSTHLCQEFPAANSVLEDEKGRKHSNCIIDLKTEAY 660
DB 601 ISTFLKMGILQEHPLPLSYSTHLCQEFPAANSVLEDEKGRKHSNCIIDLKTEAY 660
QY 661 GIGHGFGASTTFFLGLLSDGEREMENIFHCRLSQRNLMQWPSLOLLQPHSLSLH 720
DB 661 GIGHGFGASTTFFLGLLSDGEREMENIFHCRLSQRNLMQWPSLOLLQPHSLSLH 720
QY 721 CLYETRNKTFLTQWMAHPEBMCVETDMELLVCTPCIFSRHYKQLQIIBORHSTWS 780
DB 721 CLYETRNKTFLTQWMAHPEBMCVETDMELLVCTPCIFSRHYKQLQIIBORHSTWS 780
QY 781 PTMVLVFRWVPVTDVYWOILFVLKVTNRLKELDLSGNSLSHSAVSLCKTLRRPCLE 840
DB 781 PTMVLVFRWVPVTDVYWOILFVLKVTNRLKELDLSGNSLSHSAVSLCKTLRRPCLE 840
QY 841 TURLAGCGLTABDCDOLAFGLRANQTLIELDSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
DB 841 TURLAGCGLTABDCDOLAFGLRANQTLIELDSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
QY 901 LVSQGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRLLCEGRHPRACKLIRGLDQ 960
DB 901 LVSQGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRLLCEGRHPRACKLIRGLDQ 960
QY 961 TILSDMEOELRALBOEKPOLLIIFSRKPSVWTPTEGLDTGEMSNSTSLKQRLGSEBA 1020
DB 961 TILSDMEOELRALBOEKPOLLIIFSRKPSVWTPTEGLDTGEMSNSTSLKQRLGSEBA 1020
QY 1021 ASHVAQANILKLDVSKIFPIAIESSPEVVPVELCVPSASQGLHTKPLGTDDEFW 1080
DB 1021 ASHVAQANILKLDVSKIFPIAIESSPEVVPVELCVPSASQGLHTKPLGTDDEFW 1080
QY 1081 GPTGVATVEVDKKNLRYHVPVAGSYRMPNTGCLFVWRKAVTYIEFCWVDQFLG-BI 1139
DB 1081 GPTGVATVEVDKKNLRYHVPVAGSYRMPNTGCLFVWRKAVTYIEFCWVDQFLG-BI 1139
QY 1140 NPOHSMVAVGPLLDJKAEP-GAVEBAVHLPHFVALOGGVDTSLFQMAHFKKEGMLLEKPA 1198
DB 1140 NPOHSMVAVGPLLDJKAEP-GAVEBAVHLPHFVALOGGVDTSLFQMAHFKKEGMLLEKPA 1198
QY 1141 QHHEOMLVGRLPDTVTAEBEVAEIHLPFISLQ-GEVDVSWFLVAHKNMGWLEHNA 1199
DB 1141 QHHEOMLVGRLPDTVTAEBEVAEIHLPFISLQ-GEVDVSWFLVAHKNMGWLEHNA 1199
QY 1199 RVELHHIYLENDSFSLGVLKMTINLRFIPVTSVLLYHVRHEEYTPHLYLIPSDCS 1258
DB 1200 RVEPFAVLESFSLGVLKMTINLRFIPVTSVLLYHVRHEEYTPHLYLIPSDCS 1259
QY 1259 IR-----KELCYRSPGEDOLF 1276
DB 1260 LKKAIDDEDRPHGVRLOTSPMEPLNFGSSYIVANSANLKVMPBELKLSYSPEDOLF 1319
QY 1277 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDMLPATLLIPPAIIVPSPLDAPQLH 1336
DB 1320 SKFYAGQMKEPQLBITERHGTIVMDIEVKVDQIVAAAP-----FSGAA--- 1369
QY 1337 FVDQYREQLIARYTSVEVLDKLGVLSSOBYERYLAENTRPSQMKLFSLISQSWDRK 1395
DB 1370 FYKENHROLOAMWMDLQVLDLQDNEVLTENKELVEDEKTRQSKNEALSMVEKKGDL 1429
QY 1396 KDLGYOALKETHPHLIMEL 1415

```

```

Db      1430 ALDVLFRSISERDPYLSYL 1449

RESULT 12
US-09-388-221-12
; Sequence 12, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1 Card Proteins Involved in Cell Death Regu
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1424
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-388-221-12

Query Match      81.9%; Score 6174; DB 10; Length 1424;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

QY      1 MAGGAMGRILACYLEFLKKEBELKEFOLLANKAHSRSSSGETPAOPEKTSGMENVASYLVAQ 60
Db      1 MAGGAMGRILACYLEFLKKEBELKEFOLLANKAHSRSSSGETPAOPEKTSGMENVASYLVAQ 60
QY      61 YGEORANDLALHTHWEOMKLRSLCAOAOAGHSPFPSPSPHIGSOTSTAVLMPW 120
Db      61 YGEORANDLALHTHWEOMKLRSLCAOAOAGHSPFPSPSPHIGSOTSTAVLMPW 120
QY      121 IHEIPAGCTGSEKRVLRQLPDTSGRRKREISASLLYALPSPDPHESPQSPMAPTST 180
Db      121 IHEIPAGCTGSEKRVLRQLPDTSGRRKREISASLLYALPSPDPHESPQSPMAPTST 180
QY      121 IHEIPAGCTGSEKRVLRQLPDTSGRRKREISASLLYALPSPDPHESPQSPMAPTST 180
Db      121 IHEIPAGCTGSEKRVLRQLPDTSGRRKREISASLLYALPSPDPHESPQSPMAPTST 180
QY      181 AVLSGWSGPPQSLAPRBOAPGTOWPLDETSGIYYTBRERERKSEKRPMAAVVGT 240
Db      181 AVLSGWSGPPQSLAPRBOAPGTOWPLDETSGIYYTBRERERKSEKRPMAAVVGT 240
QY      241 PPOAHTSIQPHHHWEPVRESLSCTWPKNEDFNQKTOULLORPPRQODPLVKRSM 300
Db      241 PPOAHTSIQPHHHWEPVRESLSCTWPKNEDFNQKTOULLORPPRQODPLVKRSM 300
QY      241 PPOAHTSIQPHHHWEPVRESLSCTWPKNEDFNQKTOULLORPPRQODPLVKRSM 300
Db      241 PPOAHTSIQPHHHWEPVRESLSCTWPKNEDFNQKTOULLORPPRQODPLVKRSM 300
QY      301 PDVYENNGHILIRDLFGPLDTPERIVILQGAAGIGKSTLARQVKEAMGRQLYGDR 360
Db      301 PDVYENNGHILIRDLFGPLDTPERIVILQGAAGIGKSTLARQVKEAMGRQLYGDR 360
QY      361 FOHVYFSPSCRELAOSKVSLAELIGKDGTAAPAPIRQILSRPERLLFLDGVDERGWLQ 420
Db      361 FOHVYFSPSCRELAOSKVSLAELIGKDGTAAPAPIRQILSRPERLLFLDGVDERGWLQ 420
QY      361 FOHVYFSPSCRELAOSKVSLAELIGKDGTAAPAPIRQILSRPERLLFLDGVDERGWLQ 420
Db      361 FOHVYFSPSCRELAOSKVSLAELIGKDGTAAPAPIRQILSRPERLLFLDGVDERGWLQ 420
QY      421 EPSSBELCLHWSQOPADALLSLGKTLTPRASFITRTTALQNLISLEQARVETLG 480
Db      421 EPSSBELCLHWSQOPADALLSLGKTLTPRASFITRTTALQNLISLEQARVETLG 480
QY      421 EPSSBELCLHWSQOPADALLSLGKTLTPRASFITRTTALQNLISLEQARVETLG 480
Db      421 EPSSBELCLHWSQOPADALLSLGKTLTPRASFITRTTALQNLISLEQARVETLG 480
QY      481 FSESRSKRYFYRFTDERQAIRAFRLVSNKELMALCVPMVSWLACTCLMOMKREKL 540
Db      481 FSESRSKRYFYRFTDERQAIRAFRLVSNKELMALCVPMVSWLACTCLMOMKREKL 540
QY      481 FSESRSKRYFYRFTDERQAIRAFRLVSNKELMALCVPMVSWLACTCLMOMKREKL 540
Db      481 FSESRSKRYFYRFTDERQAIRAFRLVSNKELMALCVPMVSWLACTCLMOMKREKL 540
QY      541 TLTSKTTTTLCLHYLAQALQAPLQPLQRLDCLSLAABGIWQKTLFSPDDLRKGLDCAI 600
Db      541 TLTSKTTTTLCLHYLAQALQAPLQPLQRLDCLSLAABGIWQKTLFSPDDLRKGLDCAI 600
QY      541 TLTSKTTTTLCLHYLAQALQAPLQPLQRLDCLSLAABGIWQKTLFSPDDLRKGLDCAI 600
Db      541 TLTSKTTTTLCLHYLAQALQAPLQPLQRLDCLSLAABGIWQKTLFSPDDLRKGLDCAI 600
QY      601 ISTFLKMGIILOEHPRLPSYSFIHLCTQFPFAMSVLDEKGRGHSNCIIDLKTTLEAY 660
Db      601 ISTFLKMGIILOEHPRLPSYSFIHLCTQFPFAMSVLDEKGRGHSNCIIDLKTTLEAY 660
QY      601 ISTFLKMGIILOEHPRLPSYSFIHLCTQFPFAMSVLDEKGRGHSNCIIDLKTTLEAY 660
Db      601 ISTFLKMGIILOEHPRLPSYSFIHLCTQFPFAMSVLDEKGRGHSNCIIDLKTTLEAY 660
QY      661 GIHGLFGASTTRFLGLSDGEREMENIFHCRLSGGRNLQMOWPVSILQLLQPHSLSESLH 720
Db      661 GIHGLFGASTTRFLGLSDGEREMENIFHCRLSGGRNLQMOWPVSILQLLQPHSLSESLH 720

```

```

Db      661 GIHGLFGASTTRFLGLSDGEREMENIFHCRLSGGRNLQMOWPVSILQLLQPHSLSESLH 720
QY      721 CLYETRNKTFITQVMAHFEEMGCVETDMEILVCTFCIKESRHVYKQLIEGROHRTWS 780
Db      721 CLYETRNKTFITQVMAHFEEMGCVETDMEILVCTFCIKESRHVYKQLIEGROHRTWS 780
QY      781 PTMVVFRWVPVTPAVYQIILFSYKATRNKLELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
Db      781 PTMVVFRWVPVTPAVYQIILFSYKATRNKLELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
QY      841 TLRLAGGLFAEDCKDLAFGRANQTTLELDLSFNVLTDAGAKLRCORLROPSCKLQRLQ 900
Db      841 TLRLAGGLFAEDCKDLAFGRANQTTLELDLSFNVLTDAGAKLRCORLROPSCKLQRLQ 900
QY      901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLIRLGIDQ 960
Db      901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLIRLGIDQ 960
QY      901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLIRLGIDQ 960
Db      901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLIRLGIDQ 960
QY      961 TTLSDERQBLRALQEKRPOLLIFSRKRPSTVTTEGLDTGEMENSTSLKRORLGSERA 1020
Db      961 TTLSDERQBLRALQEKRPOLLIFSRKRPSTVTTEGLDTGEMENSTSLKRORLGSERA 1020
QY      958 -----KPSVMTPEGLDTGEMENSTSLKRORLGSERA 990
Db      958 -----KPSVMTPEGLDTGEMENSTSLKRORLGSERA 990
QY      1021 ASHYAQNKLKLDVSKIFPIAEIAEESSEPVVPELLCVSPASQDLPHTKPLGTDDEFW 1080
Db      991 ASHYAQNKLKLDVSKIFPIAEIAEESSEPVVPELLCVSPASQDLPHTKPLGTDDEFW 1050
QY      1081 GPTGPVATVENVKCKNLRYRHPVAGSYRMPNTGLCFVMBEAVTEIEFCVMDQFLG-BI 1139
Db      1051 GPEGNVDELIDKSTNYSVWFPTAGWYLSATGLGFLVRDEVVTIAFGSMOHLADL 1110
QY      1140 NPOHSMVAVGPILDIKAP- GAVEAVHLPHFVALQGHVDTSLFQMAHFKEGMLLEKPA 1198
Db      1111 QHBEQWLVGFLFVNTAEPEEVAEHLPHFISLQ-GEVDVSWFLVAHFNKEGAVLEHRA 1169
QY      1199 RVELHIVLENPSSPGVLLKMHNLRFIPVSVLLLYRVPVEEYTFLYLIPSDCS 1258
Db      1170 RVEEFVAVLESBSLSMGLIRASGTRLSIPITSNTLIYHPPEDIKFLYLIVPSAL 1229
QY      1259 IR-----KELCYRSPGEDOLF 1276
Db      1230 LTKAIDDEDRFHGVRLOTSPPEMLNFGSSYIVSNSANLKVMPKEKLSYRSPGEIOHF 1289
QY      1277 SEFYVGHLSGIRLQVNDKXDETLVMEALYKPGDLMATTLIPARLAVSPRLAPQLH 1336
Db      1290 SKFYAGMKRPIQLEITEKHGTLVWDTGVKPVDLQVLAASAPP-----PFGAA-- 1339
QY      1337 FVDYRQELARVTSVAVLIDKLH-GQVLSQOYERVLAETVPSOMKLFSLSQSWDRK 1395
Db      1340 FVKNHMQQLARMQDLDGVDLQDNEVLTENEKELVQEKTRQSKNEALLSMYKKGDL 1399
QY      1396 CKDGLYQALNETHPHLIMEI 1415
Db      1400 ALDVLFRSISERDPYLSYL 1419

RESULT 13
US-10-407-866-92
; Sequence 92, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: Encoding Nucleic Acids, and Methods of Use
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 764

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 2662.52 Seconds

(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-7

Perfect score: 740

Sequence: 1 cgcgcctccgcgtcgcgcggg.....atcgaagccgcttga 740

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_man:*
37: em_hcg_vrt:*
38: em_ey:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	740	100.0	740	6 AR256267	AR256267 Sequence
2	740	100.0	740	6 AX082246	AX082246 Sequence
3	740	100.0	740	6 AX536208	AX536208 Sequence
4	740	100.0	740	6 AX536223	AX536223 Sequence
5	740	100.0	740	9 AF384665	AF384665 Homo sapi
6	734	99.2	770	6 AX118619	AX118619 Homo sapi
7	734	99.2	770	6 AF184073	AF184073 Homo sapi
8	734	99.2	782	6 AX459863	AX459863 Sequence
9	734	99.2	782	6 AB023416	AB023416 Homo sapi
10	733	99.1	779	6 AX017270	AX017270 Sequence
11	733	99.1	779	6 AX524974	AX524974 Sequence
12	733	99.1	779	6 BD134441	BD134441 Human nuc
13	730.8	98.8	785	6 BD057255	BD057255 Apoptosis
14	713	96.4	740	9 AK000211	AK000211 Homo sapi
15	687	92.8	712	9 AF310103	AF310103 Homo sapi
16	680	91.9	772	9 EC013569	EC013569 Homo sapi
17	610	82.4	713	6 AX118641	AX118641 Sequence
18	610	82.4	713	9 AF255794	AF255794 Homo sapi
19	585	79.1	585	6 AR256268	AR256268 Sequence
20	585	79.1	585	6 AX082248	AX082248 Sequence
21	585	79.1	585	6 AX536210	AX536210 Sequence
22	448	60.5	450	12 BT007571	BT007571 Synthetic
23	437.6	59.1	768	9 BC004470	BC004470 Homo sapi
24	416.8	56.3	588	4 AB050006	AB050006 Bos tauri
25	378	51.1	405	6 AX118643	AX118643 Sequence
26	360	48.6	2821	6 AX118618	AX118618 Sequence
27	360	48.6	2821	9 AF184072	AF184072 Homo sapi
28	360	48.6	127769	9 AC009088	AC009088 Homo sapi
29	356	48.1	356	6 AX118626	AX118626 Sequence
30	336	45.4	732	10 BC008252	BC008252 Mus muscu
31	336	45.4	770	10 AB032249	AB032249 Mus muscu
32	336	45.4	777	6 AR256273	AR256273 Sequence
33	336	45.4	777	6 AX082258	AX082258 Sequence
34	336	45.4	777	6 AX536205	AX536205 Sequence
35	336	45.4	777	6 AX536222	AX536222 Sequence
36	336	45.4	803	6 AX118637	AX118637 Sequence
37	333	43.5	701	10 AF310104	AF310104 Mus muscu
38	322	43.5	626	6 AX118621	AX118621 Sequence
39	321	43.4	340	6 AX118622	AX118622 Sequence
40	316.2	42.7	579	6 AR256274	AR256274 Sequence
41	316.2	42.7	579	6 AX082260	AX082260 Sequence
42	316	42.7	582	10 AB053165	AB053165 Rattus no
43	309.8	41.9	321	6 AX408419	AX408419 Sequence
44	293.8	39.7	579	6 AX536207	AX536207 Sequence
45	276.8	37.4	605	6 AX118639	AX118639 Sequence

ALIGNMENTS

RESULT 1
AR256267
LOCUS AR256267 740 bp DNA
DEFINITION Sequence 48 from patent US 6482933.
ACCESSION AR256267
VERSION AR256267.1 GI:27305690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 740)
AUTHORS Bertin,J.
TITL Molecules of the card-related protein family and uses thereof
JOURNAL Patent: US 6482933-A 48 19-NOV-2002;
FEATURES Location/Qualifiers

RESULT 3	AXS36208	740 bp	DNA	linear	PAT 22-NOV-2002
LOCUS	AXS36208	740 bp	DNA	linear	PAT 22-NOV-2002
DEFINITION	Sequence 4 from M00244354.				
ACCESSION	AXS36208				
VERSION	AXS36208.1				
KEYWORDS	GI:25262599				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	Bertin, J.				
AUTHORS	Novel molecules of the card-related protein family and uses thereof				
TITLE	Patent: WO 0244354-A 4 06-JUN-2002;				
JOURNAL	MILLENNIUM PHARMACEUTICALS, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..740				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	54..641				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD57246.1"				
	/db_xref="GI:25262600"				
	/translation="MGRARDAIDLALLENLTAEELKKPKLLSVLRGYGRIRGALALSMALDITDLVSPLETTGABLTANVLRPMGLQEMAGOLQATTHGSGAAPGIGAPROGAARKGHFTIOHPALAIARNTNEMLDALYXKUTDBOYQAVRAEPTNSKKRKLFSTTPAMNWKCDLLQALRESQSIIVEDLERS"				
CDS					
BASE COUNT	146 a 238 c 236 g 120 t				
ORIGIN					
Query Match	100.0%; Score 740; DB 6; Length 740;				
Best Local Similarity	100.0%; Pred. No. 3.1e-107;				
Matches 740; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 CGCGTCGGGCTGCAGCGGGGGGTAGCGGCGCGCGCGGGGATCTTGAGCCATGGGGC	60			
DB	1 CGCGTCGGGCTGCAGCGGGGGGTAGCGGCGCGCGCGGGGATCTTGAGCCATGGGGC	60			
QY	61 GGGCGCGGAGCGCCATCTTGATGCGCTGAGAGAACTGACCGCGGAGAGCTCAAGAGT	120			
DB	61 GGGCGCGGAGCGCCATCTTGATGCGCTGAGAGAACTGACCGCGGAGAGCTCAAGAGT	120			
QY	121 TCAAGCTGAAGCTGTGTCCGGTCCGGTCCGGTCCGGTCAAGGGGCTACCGCGGGGGCG	180			
DB	121 TCAAGCTGAAGCTGTGTCCGGTCCGGTCCGGTCCGGTCAAGGGGCTACCGCGGGGGCG	180			
QY	181 CGCTGCTGTCATGAGCGCTTGAGACTCACCGACAAAGCTGTCAGCTTCACTGAGAGA	240			
DB	181 CGCTGCTGTCATGAGCGCTTGAGACTCACCGACAAAGCTGTCAGCTTCACTGAGAGA	240			
QY	241 CCTAAGGGGGCGAGCTCAACGGTTAACGTGTCGCGGAGCATGGGGCTGCAAGAGATGGCG	300			
DB	241 CCTAAGGGGGCGAGCTCAACGGTTAACGTGTCGCGGAGCATGGGGCTGCAAGAGATGGCG	300			
QY	301 GGCAGCTGAGGGCGGCGCAAGCGAGCTCTGAGAGCGCGCGAGCTGGATTCAGGGCC	360			
DB	301 GGCAGCTGAGGGCGGCGCAAGCGAGCTCTGAGAGCGCGCGAGCTGGATTCAGGGCC	360			
QY	361 CTCTCAAGTCGGACCGCAAGCGGCTTGACCTTATAGACCAAGCAAGCGGGGCTGGCTTA	420			
DB	361 CTCTCAAGTCGGACCGCAAGCGGCTTGACCTTATAGACCAAGCAAGCGGGGCTGGCTTA	420			
QY	421 TCGCAGAGGTCACAACGTTAGTGGCTGTGGAGCTCTGTAAGGGAAGGCTCTGAGAGG	480			
DB	421 TCGCAGAGGTCACAACGTTAGTGGCTGTGGAGCTCTGTAAGGGAAGGCTCTGAGAGG	480			

Db		421	TCCGGAGGAGTCACAAACGTTAGTGCGCTGCATGCTCTGTACGGAGAAGTCTTGACGG	480
Oy		481	ATGAGCAGTACCAAGCAGTGTCCGGGCCGACGCCAACCAACCAAGACAATGTGGGAAGTCT	540
Db		481	ATGAGCAGTACCAAGCAGTGTCCGGGCCGACGCCAACCAACCAAGACAATGTGGGAAGTCT	540
Oy		541	TCAATTTCACACACGACTTGAAACTGAGACTGTGCAAGACTTGCTCCTCCAGGCCCTTAAGGG	600
Db		541	TCAATTTCACACACGACTTGAAACTGAGACTGTGCAAGACTTGCTCCTCCAGGCCCTTAAGGG	600
Oy		601	AGTCCAGTCTTACCTGTGTGAGAGACCTTGAGCGGAGCTGAGAGCTCTTCCACGCAACAC	660
Db		601	AGTCCAGTCTTACCTGTGTGAGAGACCTTGAGCGGAGCTGAGAGCTCTTCCACGCAACAC	660
Oy		661	TCCGCTCAGCCCTCTGGCAATCCCACCAAATCATCTGAAATCTGATCTTTTATACAAAT	720
Db		661	TCCGCTCAGCCCTCTGGCAATCCCACCAAATCATCTGAAATCTGATCTTTTATACAAAT	720
Oy		721	ATACGAAAAGCCAGCTTGA A 740	
Db		721	ATACGAAAAGCCAGCTTGA A 740	
RESULT 4				
AXS36223/c				
LOCUS		AXS36223	740 bp	DNA linear PAT 22-NOV-2002
DEFINITION		Sequence 19 from Patent WO244354.		
ACCSSION		AXS36223		
VERSION		AXS36223.1 GI:25262610		
KEYWORDS				
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL		Bertin,J. Novel molecules of the card-related protein family and uses thereof Patent: WO 0244354-A 19 06-JUN-2002; MILENIMUM PHARMACEUTICALS, INC. (US) Location/Qualifiers		
FEATURES		1..740		
source		/organism="Homo sapiens" /mol_type="genomic DNA"		
		/db_xref="taxon:9606"		
		/db_xref="taxon:9606"		
BASE COUNT		120 a 236 c 238 g 146 t		
ORIGIN				
		Query Match 100.0%; Score 740; DB 6; Length 740;		
		Best Local Similarity 100.0%; Pred. No.3.1e-107; Mismatches 0; Indels 0; Gaps 0;		
		Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Oy		1	C GCCTCCGCGCTGCACGCGGGGTGAGCGCGCGACGCGCGCGGGATCTTGAGCCATGSGGC	60
Db		740	C GCCTCCGCGCTGCACGCGGGGTGAGCGCGCGCGCGCGGGATCTTGAGCCATGSGGC	681
Oy		61	G CGCGCGCGAGCCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGAAGT	120
Db		680	G CGCGCGCGAGCCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGAAGT	621
Oy		121	T CAAGCTGAAGCTGCTGTCCGCTGCGCTGCGCGAGGAGCTACGGGCGATCCCGCGGGGCG	180
Db		620	T CAAGCTGAAGCTGCTGTCCGCTGCGCTGCGCGAGGAGCTACGGGCGATCCCGCGGGGCG	561
Oy		181	C GCTGCTGTCCATGAGACGCTTGACCTCAACCGAACAAGTGTACGCTTACCTGAGAGA	240
Db		560	C GCTGCTGTCCATGAGACGCTTGACCTCAACCGAACAAGTGTACGCTTACCTGAGAGA	501
Oy		241	C CTACGGGCGCGAGCTCACCGCTAAAGTGTGCGCGACATGAGGCTGTGACAGAGATGCGG	300
Db		500	C CTACGGGCGCGAGCTCACCGCTAAAGTGTGCGCGACATGAGGCTGTGACAGAGATGCGG	441
Oy		301	G G C A G C T G C A G G C G G C C A C G A C C A G G G C T C T G A G C C G G C C A G C T G G A A T C C A G G C C	360

Db 440 GGCAGCTGACGAGGCGACACGAGGCTCTGAGGCCGCGCAGCTGGATCCAGGCC 381
 QY 361 CTCCTCAGTGGGACGACGAGGCGCTTGAATATAGACGACGAGCGGCTGCGCTTA 420
 Db 380 CTCCTCAGTGGGACGACGAGGCGCTTGAATATAGACGACGAGCGGCTGCGCTTA 321
 QY 421 TCCGAGAGGCTCACAAAGTTGAGTGGCTGAGTGTCTGTACGGGAAGTCTTGA 480
 Db 320 TCCGAGAGGCTCACAAAGTTGAGTGGCTGAGTGTCTGTACGGGAAGTCTTGA 261
 QY 481 ATGAGCAGTACGAGGAGTGGGCGGAGCCGACCAACCAAGCAAGATCGGAGCTCT 540
 Db 260 ATGAGCAGTACGAGGAGTGGGCGGAGCCGACCAACCAAGCAAGATCGGAGCTCT 201
 QY 541 TCAGTTTCAACACGAGCTTGAATCTGACCTGCAAGAGTCTTCTTCCAGGCTTAAGG 600
 Db 200 TCAGTTTCAACACGAGCTTGAATCTGACCTGCAAGAGTCTTCTTCCAGGCTTAAGG 141
 QY 601 AGTCCAGTCTTACCTGTGAGAGGAGCTGAGAGGAGCTGAGGCTCTTCCAGCAAC 660
 Db 140 AGTCCAGTCTTACCTGTGAGAGGAGCTGAGAGGAGCTGAGGCTCTTCCAGCAAC 81
 QY 661 TCCGTCAGGCGCTTGGCAATCCGCAATCATCTGTAATCTGATCTTTTATACCAAT 720
 Db 80 TCCGTCAGGCGCTTGGCAATCCGCAATCATCTGTAATCTGATCTTTTATACCAAT 21
 QY 721 ATACGAAAGCCAGCTTGAA 740
 Db 20 ATACGAAAGCCAGCTTGAA 1

RESULT 5
 AF384665 740 bp mRNA linear PRI 19-JUN-2001
 LOCUS Homo sapiens caspase recruitment domain protein 5 mRNA, complete
 DEFINITION cde.
 ACCESSION AF384665
 VERSION AF384665.1 GI:14488058
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 740)
 AUTHORS Bertin,J.
 TITLE CARD5 Protein is a CARD/PYRIN family member that is involved in
 CARD domain signal transduction
 JOURNAL CARD5 Protein is a CARD/PYRIN family member that is involved in
 CARD domain signal transduction
 REFERENCE 2 (bases 1 to 740)
 AUTHORS Bertin,J.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals
 Inc., 640 Memorial Drive, Cambridge, MA 02139, USA
 FEATURES
 source
 1..740
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 54..641
 /note="CARD5, CARD/PYRIN family member; bipartite protein
 comprised of an N-terminal PYRIN domain and a C-terminal
 CARD domain; involved in apoptosis signal transduction"
 /codon_start=1
 /product="caspase recruitment domain protein 5"
 /protein_id="AAK63850.1"
 /db_xref="GI:14488058"
 /translation="MGRARDALIDALENTTAEELKKFKLKLKLVPLREGVGRIPRGAL
 LMSDALDLDKLVSYFLTYGAEITANVLRDMLQDMAGOLAAATGSGGAPAGIOA
 POSAKRGRLHIDQHRALLIARVTVMELDLKXKVLTDDEYQAVRAEPTNPSMR
 KLFSTPDANMWTCKDLQLRKESQSYLVEDERS"

CDS
 54..641
 /note="CARD5, CARD/PYRIN family member; bipartite protein
 comprised of an N-terminal PYRIN domain and a C-terminal
 CARD domain; involved in apoptosis signal transduction"
 /codon_start=1
 /product="caspase recruitment domain protein 5"
 /protein_id="AAK63850.1"
 /db_xref="GI:14488058"
 /translation="MGRARDALIDALENTTAEELKKFKLKLKLVPLREGVGRIPRGAL
 LMSDALDLDKLVSYFLTYGAEITANVLRDMLQDMAGOLAAATGSGGAPAGIOA
 POSAKRGRLHIDQHRALLIARVTVMELDLKXKVLTDDEYQAVRAEPTNPSMR
 KLFSTPDANMWTCKDLQLRKESQSYLVEDERS"
 BASE COUNT 146 a 238 c 236 g 120 t
 ORIGIN

Query Match 100.0%; Score 740; DB 9; Length 740;
 Best Local Similarity 100.0%; Pred. No. 3.1e-107;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTCGCGGTGACGAGGAGGTGAGCGGCGGAGCGGCGGAGTCTTGAGGCAATGAGG 60
 Db 1 CGGTCGCGGTGACGAGGAGGTGAGCGGCGGAGCGGCGGAGTCTTGAGGCAATGAGG 60
 QY 61 GCGGCGGCGAGCGGCATCTGATGCGCTGAGAGAACTTGAACCGCGAGAGACTCAAGAGT 120
 Db 61 GCGGCGGCGAGCGGCATCTGATGCGCTGAGAGAACTTGAACCGCGAGAGACTCAAGAGT 120
 QY 121 TCAAGCTGAAGCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 Db 121 TCAAGCTGAAGCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 181 CGCTGTCTGTCATGAGAGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 240
 Db 181 CGCTGTCTGTCATGAGAGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 240
 QY 241 CCTACGCGCGCGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCT 300
 Db 241 CCTACGCGCGCGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCT 300
 QY 301 GCGAGCTGACGAGCGGCGACGACGAGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGG 360
 Db 301 GCGAGCTGACGAGCGGCGACGACGAGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGG 360
 QY 361 CTCCTCAGTCCGAGGCGACGAGGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 420
 Db 361 CTCCTCAGTCCGAGGCGACGAGGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 420
 QY 421 TCCGAGAGGCTCACAAAGTTGAGTGGCTGAGTGTCTGTACGGGAAGTCTTGAAGAG 480
 Db 421 TCCGAGAGGCTCACAAAGTTGAGTGGCTGAGTGTCTGTACGGGAAGTCTTGAAGAG 480
 QY 481 ATGAGCAGTACGAGGAGTGGGCGGAGCCGACCAACCAAGCAAGATCGGAGCTCT 540
 Db 481 ATGAGCAGTACGAGGAGTGGGCGGAGCCGACCAACCAAGCAAGATCGGAGCTCT 540
 QY 541 TCAGTTTCAACACGAGCTTGAATCTGACCTGCAAGAGTCTTCTTCCAGGCTTAAGG 600
 Db 541 TCAGTTTCAACACGAGCTTGAATCTGACCTGCAAGAGTCTTCTTCCAGGCTTAAGG 600
 QY 601 AGTCCAGTCTTACCTGTGAGAGGAGCTGAGAGGAGCTGAGGCTCTTCCAGCAAC 660
 Db 601 AGTCCAGTCTTACCTGTGAGAGGAGCTGAGAGGAGCTGAGGCTCTTCCAGCAAC 660
 QY 661 TCCGTCAGGCGCTTGGCAATCCGCAATCATCTGTAATCTGATCTTTTATACCAAT 720
 Db 661 TCCGTCAGGCGCTTGGCAATCCGCAATCATCTGTAATCTGATCTTTTATACCAAT 720
 QY 721 ATACGAAAGCCAGCTTGAA 740
 Db 721 ATACGAAAGCCAGCTTGAA 740

RESULT 6
 AX118619 770 bp DNA linear PAT 11-MAY-2001
 LOCUS Sequence 2 from Patent WO0129235.
 DEFINITION AX118619
 ACCESSION AX118619.1 GI:14035570
 VERSION AX118619.1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Vertino,P.M.
 TITLE Tms1 compositions and methods of use
 JOURNAL Patent: WO 0129235-A 2 26-APR-2001;
 Emory University (US)

QY 127 TGAAGCTGCTGTGCGGCGCTGCGAGGAGGCTACCGGCGCATCCGCGGCGGCGCTGC 186
 DB 148 TGAAGCTGCTGTGCGGCGCTGCGAGGAGGCTACCGGCGCATCCGCGGCGGCGCTGC 207
 QY 187 TGTTCATGAGCGGCTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 246
 DB 208 TGTTCATGAGCGGCTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 267
 QY 247 GGGCCGAGCTCAACCGCTTAACTGTCTGCGGAGCATGAGGCTTCAAGCTTCAAGCT 306
 DB 268 GGGCCGAGCTCAACCGCTTAACTGTCTGCGGAGCATGAGGCTTCAAGCTTCAAGCT 327
 QY 307 TGAAGCGGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 366
 DB 328 TGAAGCGGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 387
 QY 367 AGTCGCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 426
 DB 388 AGTCGCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 447
 QY 427 GGGTCACAAACGTTGAGT 486
 DB 448 GGGTCACAAACGTTGAGT 507
 QY 487 AGTACCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 546
 DB 508 AGTACCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 567
 QY 547 TCACACGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 606
 DB 568 TCACACGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 627
 QY 607 AGTCTCACTGT 666
 DB 628 AGTCTCACTGT 687
 QY 667 CAGCCCTGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 726
 DB 688 CAGCCCTGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 747
 QY 727 AAAGCCAGCTTGAA 740
 DB 748 AAAGCCAGCTTGAA 761
 RESULT 8
 AX459863 782 bp DNA linear PAT 08-JUL-2002
 LOCUS Sequence 6 from Patent WO0240668.
 DEFINITION AX459863
 ACCESSION AX459863
 VERSION AX459863.1 GI:21725640
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1. 782
 AUTHORS Taniguchi, J. and Martinon, F.
 TITLE Proteins and dna sequences underlying these proteins used for
 JOURNAL treating inflammations
 Apotech Research and Development Ltd. (CH)
 FEATURES
 source
 1. 782
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="PyCard.cdna"
 BASE COUNT 158 a 251 c 246 g 127 t
 ORIGIN
 Query Match 99.2%; Score 734; DB 6; Length 782;
 Best Local Similarity 100.0%; Pred. No. 2.8e-106;

Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 CCGCTGACAGCGGAGTGAAGCGGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAG 66
 DB 40 CCGCTGACAGCGGAGTGAAGCGGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAG 99
 QY 67 GGGAGCGGAGTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAG 126
 DB 100 GGGAGCGGAGTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAG 159
 QY 127 TGAAGCTGCTGTGCGGCGCTGCGAGGAGGCTACCGGCGCATCCGCGGCGGCGCTGC 186
 DB 160 TGAAGCTGCTGTGCGGCGCTGCGAGGAGGCTACCGGCGCATCCGCGGCGGCGCTGC 219
 QY 187 TGTTCATGAGCGGCTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 246
 DB 220 TGTTCATGAGCGGCTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 279
 QY 247 GGGCCGAGCTCAACCGCTTAACTGTCTGCGGAGCATGAGGCTTCAAGCTTCAAGCT 306
 DB 280 GGGCCGAGCTCAACCGCTTAACTGTCTGCGGAGCATGAGGCTTCAAGCTTCAAGCT 339
 QY 307 TGAAGCGGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 366
 DB 340 TGAAGCGGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 399
 QY 367 AGTCGCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 426
 DB 400 AGTCGCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 459
 QY 427 GGGTCACAAACGTTGAGT 486
 DB 460 GGGTCACAAACGTTGAGT 519
 QY 487 AGTACCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 546
 DB 520 AGTACCGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 579
 QY 547 TCACACGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 606
 DB 580 TCACACGAGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 639
 QY 607 AGTCTCACTGT 666
 DB 640 AGTCTCACTGT 699
 QY 667 CAGCCCTGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 726
 DB 700 CAGCCCTGCGGAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCTTGAAGCGGAGCT 759
 QY 727 AAAGCCAGCTTGAA 740
 DB 760 AAAGCCAGCTTGAA 773
 RESULT 9
 AB023416 782 bp mRNA linear PRI 18-OCT-2000
 LOCUS Homo sapiens ASC mRNA for apoptosis-associated speck-like protein
 DEFINITION containing a CARD, complete cds.
 ACCESSION AB023416
 VERSION AB023416.2 GI:10801601
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1. (sites)
 AUTHORS Masumoto, J., Taniguchi, S., Ayukawa, K., Servotham, H., Kishino, T.,
 TITLE Nilkawa, N., Hidaka, E., Katsuyama, T., Higuchi, T. and Sagata, J.
 JOURNAL ASC, a novel 22-kDa protein, aggregates during apoptosis of human
 promyelocytic leukemia HL-60 cells
 J. Biol. Chem. 274 (48), 33835-33838 (1999)

MEDLINE 20036508
 PUBMED 10567338
 REFERENCE 2 (bases 1 to 782)
 AUTHORS Maemoto, J., Sagara, J. and Taniguchi, S.
 TITLE Direct Submission
 JOURNAL Submitted (04-FEB-1999) Junya Maemoto, Shinshu University School of Medicine, Research Center on Aging and Adaptation, Aichi 3-1-1, Matsumoto, Nagano 390-8621, Japan
 (E-mail: maemoto@sch.medic.shinshu-u.ac.jp, Tel: 81-263-37-2723, Fax: 81-263-37-2724)
 COMMENT On Oct 14, 2000 this sequence version replaced gi:6482371.
 FEATURES
 source
 1..782
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p11.2-12"
 1..782
 /gene="ASC"
 87..674
 /gene="ASC"
 /codon_start=1
 /product="apoptosis-associated speck-like protein containing a CARD"
 /protein_id="BA87339.2"
 /db_xref="GI:10801602"
 /translation="MGKADPDLIDALENLTAELKKEFKLLSVPLREGYRIPRGAL
 LSNDAIDTDLVLFYLETGALFETANVLRDMGQEMAGQLOAATHGSGAAGGLOA
 PPSAAKGGHFDHRAALFARLVNTWMLDALYGVULDEQVQARABETNSKMR
 KLSFTYAMWNTCDLILQALRESQSLVEDLERS"
 BASE COUNT 158 a 251 c 246 g 127 t
 ORIGIN
 Query Match 99.2%; Score 734; DB 9; Length 782;
 Best Local Similarity 100.0%; Pred. No. 2.8e-106; Indels 0; Gaps 0;
 Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 CGGCTGACGCGGGGTGAGCGGCGGCGGCGGCGGATCTTGAGCCATGCGGCGCGCGC 66
 Db 40 CGGCTGACGCGGGGTGAGCGGCGGCGGCGGCGGATCTTGAGCCATGCGGCGCGCGC 99
 QY 67 GCGACGCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGATTCAAGC 126
 Db 100 GCGACGCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGATTCAAGC 159
 QY 127 TGAAGCTGCTCGGTGCGCTGCGGCGGCGGCTACCGGCGGCGGCGGCTGC 186
 Db 160 TGAAGCTGCTCGGTGCGCTGCGGCGGCGGCTACCGGCGGCGGCGGCTGC 219
 QY 187 TGTCCATGAGCGCTTGACCTCAACGACAGCTGATCAGCTTACCTGAGACTAAG 246
 Db 220 TGTCCATGAGCGCTTGACCTCAACGACAGCTGATCAGCTTACCTGAGACTAAG 279
 QY 247 GCGCGGAGCTACCGCTTACCTGCTGCGGCGGCGGCTGCGGAGATGCGCGGCGGCGC 306
 Db 280 GCGCGGAGCTACCGCTTACCTGCTGCGGCGGCGGCTGCGGAGATGCGCGGCGGCGC 339
 QY 307 TGCAGGGGGCGGCGGCGGCTGAGCGGCGGCGGCGGCTGAGATCCAGGCGCCTCTC 366
 Db 340 TGCAGGGGGCGGCGGCGGCTGAGCGGCGGCGGCGGCTGAGATCCAGGCGCCTCTC 399
 QY 367 AGTCGAGGCGGCGGCGGCTGACCTTATAGACAGACCGGAGCTGCGCTTATCGGA 426
 Db 400 AGTCGAGGCGGCGGCGGCTGACCTTATAGACAGACCGGAGCTGCGCTTATCGGA 459
 QY 427 GGGTCACAAAGCTTGAAGTGTGCTGATGCTGTATAGGGAAGTCTTGAAGGATGAGC 486
 Db 460 GGGTCACAAAGTGAAGTGTGCTGATGCTGTATAGGGAAGTCTTGAAGGATGAGC 519
 QY 487 AGTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 546

Db 520 AGTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 579
 QY 547 TCACACCAAGCTTGGAATCTGAGCTGCAAGACTTGTCTTCAGGCGCTTAAGGAGTCCC 606
 Db 580 TCACACCAAGCTTGGAATCTGAGCTGCAAGACTTGTCTTCAGGCGCTTAAGGAGTCCC 639
 QY 607 AGTCTACCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
 Db 640 AGTCTACCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699
 QY 667 CAGCCCTGCGCAATCCACCAATCATCTGATCTTTTATACAAATATACGA 726
 Db 700 CAGCCCTGCGCAATCCACCAATCATCTGATCTTTTATACAAATATACGA 759
 QY 727 AAAGCCAGCTTGAA 740
 Db 760 AAAGCCAGCTTGAA 773
 RESULT 10
 AX017270 779 bp DNA linear PAT 07-SEP-2000
 LOCUS AX017270
 DEFINITION Sequence 21 from Patent WO9447669.
 ACCESSION AX017270
 VERSION AX017270.1 GI:10042188
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and
 Pilsarsky, C.
 TITLE Human nucleic acid sequences from tissue of breast tumors
 JOURNAL Patent: WO 9947669-A 21 23-SEP-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METABEN GES FUER GENOMFORSCHUN
 (DE); PILSARSKY CHRISTIAN (DE)
 FEATURES
 source
 1..779
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 149 a 252 c 251 g 127 t
 ORIGIN
 Query Match 99.1%; Score 733; DB 6; Length 779;
 Best Local Similarity 100.0%; Pred. No. 4e-106; Indels 0; Gaps 0;
 Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 CGGCTGACGCGGGGTGAGCGGCGGCGGCGGCGGATCTTGAGCCATGCGGCGCGCGC 66
 Db 47 CGGCTGACGCGGGGTGAGCGGCGGCGGCGGCGGATCTTGAGCCATGCGGCGCGCGC 106
 QY 67 GCGACGCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGATTCAAGC 126
 Db 107 GCGACGCCATCTTGATGCGCTGAGAACTTGACCGCGGAGAGCTCAAGATTCAAGC 166
 QY 127 TGAAGCTGCTCGGTGCGCTGCGGCGGCGGCTACCGGCGGCGGCGGCTGC 186
 Db 167 TGAAGCTGCTCGGTGCGCTGCGGCGGCGGCTACCGGCGGCGGCGGCTGC 226
 QY 187 TGTCCATGAGCGCTTGACCTCAACGACAGCTGATCAGCTTACCTGAGACTAAG 246
 Db 227 TGTCCATGAGCGCTTGACCTCAACGACAGCTGATCAGCTTACCTGAGACTAAG 286
 QY 247 GCGCGGAGCTACCGCTTACCTGCTGCGGCGGCGGCTGCGGAGATGCGCGGCGGCGC 306
 Db 287 GCGCGGAGCTACCGCTTACCTGCTGCGGCGGCGGCTGCGGAGATGCGCGGCGGCGC 346
 QY 307 TGCAGGGGGCGGCGGCGGCTGAGCGGCGGCGGCGGCTGAGATCCAGGCGCCTCTC 366
 Db 347 TGCAGGGGGCGGCGGCGGCTGAGCGGCGGCGGCGGCTGAGATCCAGGCGCCTCTC 406

OY		367	AGTCGGCAGCCAAAGCCAGACCGCTGCACCTTATAGACACGACCAGGCTTGATTACGGA	426
Dd		407	AGTCGGCAGCCAAAGCCAGACCGCTGCACCTTATAGACACGACCAGGCTTGATTACGGA	466
OY		427	GGGTCACAAACGTTGAGTGCTGTCTGATGTCTCTGTACGGGAAGGTCCTGACGATGAGC	486
Dd		467	GGGTTCACAAACGTTGAGTGCTGTCTGATGTCTCTGTACGGGAAGGTCCTGACGATGAGC	526
OY		487	AGTACCAAGGCAAGTGGGGGCCGACCCACCAACCCAAAGAAATGGGGAAGCTCTTCAATT	546
Dd		527	AGTACCAAGGCAAGTGGGGGCCGACCCACCAACCCAAAGAAATGGGGAAGCTCTTCAATT	586
OY		547	TCAACACACGCTTGAACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACT	606
Dd		587	TCAACACACGCTTGAACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACT	646
OY		607	AGTCTTAACCTGTGTGAGGAGCCTGAGCGGAGCTGAGGCTCTTCCAGCAACAATCCGGT	666
Dd		647	AGTCTTAACCTGTGTGAGGAGCCTGAGCGGAGCTGAGGCTCTTCCAGCAACAATCCGGT	706
OY		667	CAGCCCCCTGGAAATCCCAACCAATCATCTGAAATCTGAAATCTGAAATCTGAAATCTGA	726
Dd		707	CAGCCCCCTGGAAATCCCAACCAATCATCTGAAATCTGAAATCTGAAATCTGAAATCTGA	766
OY		727	AAAGCCAGCTTGA 739	
Dd		767	AAAGCCAGCTTGA 779	
RESULT 11				
AX524974				
LOCUS	AX524974	779 bp	DNA	linear PAT 21-NOV-2002
DEFINITION	Sequence 21 from Patent EP1236799.			
ACCESSION	AX524974			
VERSION	AX524974.1	GI:25170056		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	Specht,T., Hinemann,B., Schmitt,A., Pilarsky,C., Dahl,E. and Rosenthal,A. Human nucleic acid sequences derived from breast tumor tissue Patent: EP 1236799-A 21 04-SEP-2002;			
JOURNAL	metagen Pharmaceuticals GmbH (DE) Location/Qualifiers			
FEATURES	source 1..779 /organism="Homo sapiens" /mol_type="Genomic DNA" /db_xref="taxon:9606"			
BASE COUNT	149 a 252 c 251 g 127 t			
ORIGIN				
Query Match	99.1%; Score 733; DB 6; Length 779;			
Best Local Similarity	100.0%; Pred. No. 4e-106;			
Matches 733; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
OY		7	CGGCTGCAGCGGGGGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	66
Dd		47	CGGCTGCAGCGGGGGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	106
OY		67	GCGAGCGCATCTGGATCGCTGTGAGAACCTGACCGCGGAGAGCTCAAGAAGTTCAAGC	126
Dd		107	GCGAGCGCATCTGGATCGCTGTGAGAACCTGACCGCGGAGAGCTCAAGAAGTTCAAGC	166
OY		127	TGAAGCTGCTGTGGTGCGCTGTGCGAGAGGCTACGGGCGCATCCCGCGGGGCGGCGTGC	186
Dd		167	TGAAGCTGCTGTGGTGCGCTGTGCGAGAGGCTACGGGCGCATCCCGCGGGGCGGCGTGC	226
OY		187	TGTCCATGGAAGCCTTGAACTCACACCAAGCTGTGACGTTCTAAGTGAAGACTTACG	246

Db	227	TGTCATGAGAACGCGCTTGGAACCTCACCGACAAGAGTGTCAGCTTCTACTGAGAACCTACG	286
Qy	247	GGCGGAGGCTCACCGGCTTAAAGTGTGGGCCACATGGGCTCTGCAGAGATAGTGGCGGGCAGC	306
Db	287	GGCGGAGGCTCACCGGCTTAAAGTGTGGGCCACATGGGCTCTGCAGAGATAGTGGCGGGCAGC	346
Qy	307	TGCAGAGGCGCACGACCAAGGAGCTCTGAGAGCGCGCAGCTGGAGATCCAGGCGCCCTCTC	366
Db	347	TGCAGAGGCGCACGACCAAGGAGCTCTGAGAGCGCGCAGCTGGAGATCCAGGCGCCCTCTC	406
Qy	367	AGTGGGAGGCCAAGCCAGGCTTGCACTTTATATAGACAGCACCGGGCTGCGCTTATCGCGA	426
Db	407	AGTGGGAGGCCAAGCCAGGCTTGCACTTTATATAGACAGCACCGGGCTGCGCTTATCGCGA	466
Qy	427	GGGTCACAAAAGTTGAATGGCTGTGGATAGTGTCTGTACGGGAAGTCTCTACGGATGAGC	486
Db	467	GGGTCACAAAAGTTGAATGGCTGTGGATAGTGTCTGTACGGGAAGTCTCTACGGATGAGC	526
Qy	487	AGTACCAAGGACAGTGGGGGCGAGGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	546
Db	527	AGTACCAAGGACAGTGGGGGCGAGGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	586
Qy	547	TCAACCAAGGCTGGAACCTGGAACCTGGAAGGACTTGTCTCCAGGCGCTTACGAGAGTCC	606
Db	587	TCAACCAAGGCTGGAACCTGGAACCTGGAAGGACTTGTCTCCAGGCGCTTACGAGAGTCC	646
Qy	607	AGTCTCACTGTGTGAGAGACCTGAGAGCGAGCTGAGGCTCCTTCCAGCAACACTCCGAT	666
Db	647	AGTCTCACTGTGTGAGAGACCTGAGAGCGAGCTGAGGCTCCTTCCAGCAACACTCCGAT	706
Qy	667	CAGCCCTCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATATAGA	726
Db	707	CAGCCCTCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATATAGA	766
Qy	727	AAAGCCAGCTTGA 739	
Db	767	AAAGCCAGCTTGA 779	
RESULT 12			
BD134441		779 bp DNA linear	PAT 18-SEP-2002
LOCUS			
DEFINITION		Human nucleic acid sequence originating in mammary tumor tissue.	
ACCESSION		BD134441	
VERSION		BD134441.1	GI:23229386
KEYWORDS		JP 2002506643-A/19.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		Spefic,T., Hntzman,B., Armin,S., Pitaraki,C., Edgar,D. and Rosenthal,A.	
TITLE		Human nucleic acid sequence originating in mammary tumor tissue	
JOURNAL		Patent: JP 2002506643-A 19 05-MAR-2002;	
COMMENT		METAGEN GSELSCHAFT FUER GENOME FORSCHUNG MBH	
	OS	Homo sapiens (human)	
	PN	JP 2002506643-A/19	
	PD	05-MAR-2002	
	PF	19-MAR-1999 JP 2000536852	
	PR	20-MAR-1998 DE 198 13 839.3	
	PI	THOMAS SPEFT, BERND HINTZMAN, SHOMITT ARMIN, CHRISTIAN PIRASKI,	
	PI	DUHL EDGAR,	
	PI	ANDRE ROSENTHAL,	
	PC	C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P35/00, C07K14/47,	
	PC	C07K16/18,	
	PC	C12N1/19, C12N5/10, C12N15/00, A61K37/02, C12N5/00 CC	Human
		nucleic acid sequence originating in mammary tumor CC	
	FT	Key	Location/Qualifiers
	FT	source	1..779
FEATURES		Location/Qualifiers	/'Organism='Homo sapiens (human)''.

source 1. .779
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 149 a 252 c 251 g 127 t

ORIGIN

Query Match 99.1%; Score 733; DB 6; Length 779;
Best Local Similarity 100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGACGCGGGTGAAGCGGCGGCGGCGGCGGATCTGAGCCATGAGGCGCGCGC 66
Db 47 CGGCTGACGCGGGGTGAAGCGGCGGCGGCGGCGGATCTGAGCCATGAGGCGCGCGC 106

QY 67 GCGAGCCCATCTTGAGATCGCTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGC 126
Db 107 GCGAGCCCATCTTGAGATCGCTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGC 166

QY 127 TGAAGCTGCTGCGGTGCGGCTGCGGCGGAGGAGTACCGGCGGCGCGCTGC 186
Db 167 TGAAGCTGCTGCGGTGCGGCTGCGGCGGAGGAGTACCGGCGGCGCGCTGC 226

QY 187 TGTCCATGAGCGCTTGAGACCTCAACGAGCTGTGAGCTTCTACCTTGAGAGCTTACG 246
Db 227 TGTCCATGAGCGCTTGAGACCTCAACGAGCTGTGAGCTTCTACCTTGAGAGCTTACG 286

QY 247 GCGCCGAGCTCAACGCTTAACTGCTGCGGAGCATGAGGCTGCGAGAGATGCGCGGAGC 306
Db 287 GCGCCGAGCTCAACGCTTAACTGCTGCGGAGCATGAGGCTGCGAGAGATGCGCGGAGC 346

QY 307 TGCAGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 366
Db 347 TGCAGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 406

QY 367 AGTGGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 426
Db 407 AGTGGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 466

QY 427 GGGTCAAAAGTTGAGTGGCTGCTGAGATGCTGTGATCGGAGAGTCTTACGAGATGAGC 486
Db 467 GGGTCAAAAGTTGAGTGGCTGCTGAGATGCTGTGATCGGAGAGTCTTACGAGATGAGC 526

QY 487 AGTACCGAGGAGTGGCGGCGGCGGCGGCGGCGGAGTCCAGGCGCTCTCTCTC 546
Db 527 AGTACCGAGGAGTGGCGGCGGCGGCGGCGGCGGAGTCCAGGCGCTCTCTCTC 586

QY 547 TCACACGAGCTGGAATGAGACTGAGAGACTTGTCTCTCAAGGCGCTTAAAGGAGTCC 606
Db 587 TCACACGAGCTGGAATGAGACTGAGAGACTTGTCTCTCAAGGCGCTTAAAGGAGTCC 646

QY 607 AGTCCATCTGTGAGAGAGCTTGAAGCGGAGCTGAGAGTCTCTTCCAGCAACACTCCGGT 666
Db 647 AGTCCATCTGTGAGAGAGCTTGAAGCGGAGCTGAGAGTCTCTTCCAGCAACACTCCGGT 706

QY 667 CAGGCGGCGGAGTCCGAGCAATCATCTGATCTCTTTTATACAAATATATAGA 726
Db 707 CAGGCGGCGGAGTCCGAGCAATCATCTGATCTCTTTTATACAAATATATAGA 766

QY 727 AAAGCAGAGCTTGA 739
Db 767 AAAGCAGAGCTTGA 779

RESULT 13
BD057255 785 bp DNA linear PAT 27-AUG-2002
LOCUS BD057255
DEFINITION Apoptosis-related protein, its antibody and its DNA.
ACCESSION BD057255
VERSION BD057255.1 GI:22602861
KEYWORDS JP 2001275681-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source 1. .785
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 160 a 252 c 247 g 126 t

ORIGIN

Query Match 98.8%; Score 730.8; DB 6; Length 785;
Best Local Similarity 99.7%; Pred. No. 8.8e-106;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 CGGCTGACGCGGGTGAAGCGGCGGCGGCGGCGGATCTGAGCCATGAGGCGCGCGC 66
Db 40 CGGCTGACGCGGGTGAAGCGGCGGCGGCGGCGGATCTGAGCCATGAGGCGCGCGC 99

QY 67 GCGAGCCCATCTTGAGATCGCTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGC 126
Db 100 GCGAGCCCATCTTGAGATCGCTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGC 159

QY 127 TGAAGCTGCTGCGGTGCGGCTGCGGCGGAGGAGTACCGGCGGCGCGCTGC 186
Db 160 TGAAGCTGCTGCGGTGCGGCTGCGGCGGAGGAGTACCGGCGGCGCGCTGC 219

QY 187 TGTCCATGAGCGCTTGAGACCTCAACGAGCTGTGAGCTTCTACCTTGAGAGCTTACG 246
Db 220 TGTCCATGAGCGCTTGAGACCTCAACGAGCTGTGAGCTTCTACCTTGAGAGCTTACG 279

QY 247 GCGCCGAGCTCAACGCTTAACTGCTGCGGAGCATGAGGCTGCGAGAGATGCGCGGAGC 306
Db 280 GCGCCGAGCTCAACGCTTAACTGCTGCGGAGCATGAGGCTGCGAGAGATGCGCGGAGC 339

QY 307 TGCAGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 366
Db 340 TGCAGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 399

QY 367 AGTGGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 426
Db 400 AGTGGGCGGCGCAACGAGCGGCTTGAAGCGGCGGCGGAGTCCAGGCGCTCTCTC 459

QY 427 GGGTCAAAAGTTGAGTGGCTGCTGAGATGCTGTGATCGGAGAGTCTTACGAGATGAGC 486
Db 460 GGGTCAAAAGTTGAGTGGCTGCTGAGATGCTGTGATCGGAGAGTCTTACGAGATGAGC 519

QY 487 AGTACCGAGGAGTGGCGGCGGCGGCGGCGGCGGAGTCCAGGCGCTCTCTCTC 546
Db 520 AGTACCGAGGAGTGGCGGCGGCGGCGGCGGCGGAGTCCAGGCGCTCTCTCTC 579

QY 547 TCACACGAGCTGGAATGAGACTGAGAGACTTGTCTCTCAAGGCGCTTAAAGGAGTCC 606
Db 580 TCACACGAGCTGGAATGAGACTGAGAGACTTGTCTCTCAAGGCGCTTAAAGGAGTCC 639

QY 607 AGTCCATCTGTGAGAGAGCTTGAAGCGGAGCTGAGAGTCTCTTCCAGCAACACTCCGGT 666
Db 640 AGTCCATCTGTGAGAGAGCTTGAAGCGGAGCTGAGAGTCTCTTCCAGCAACACTCCGGT 699

QY 667 CAGGCGGCGGAGTCCGAGCAATCATCTGATCTCTTTTATACAAATATATAGA 726
Db 707 CAGGCGGCGGAGTCCGAGCAATCATCTGATCTCTTTTATACAAATATATAGA 766

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 785)
Masumoto, J., Sagara, J. and Taniguchi, S.
Apoptosis-related protein, its antibody and its DNA
Patent: JP 2001275681-A 1 09-OCT-2001;
JUNYA MASUMOTO, JUNJI AIRA, SHUNICHIRO TANIGUCHI, MEDICAL &
BIOLOGICAL LABORATORIES CO LTD
OS Homo sapiens (human)
PN JP 2001275681-A/1
PD 09-OCT-2001
PF 31-MAR-2000 JP 2000098204
PI JUNYA MASUMOTO, JUNJI SAGARA, SHUNICHIRO TANIGUCHI PC
CI2N15/09, C07K14/82, C07K16/32//CI2P21/02, CI2P21/08, CI2N15/00 CC
FH Key Location/Qualifiers.

DB	700	CAGCCCTGGCAATCCCAAAATCATCTGATCTTTTATACAAATATACA	755
Y	727	AAAGCAGCTTGA	740
DB	760	AAAGCCAGCTTGA	773
RESULT 14	AK000211	740 bp	mRNA linear
LOCUS	AK000211	740 bp	mRNA linear
DEFINITION	Homo sapiens cDNA FLJ20204 fis, clone COLF1505.		PRI 22-FEB-2000
ACCESSION	AK000211		
VERSION	AK000211.1		GI:7020146
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (sites)		
TITLE	Watanabe, K., Kunagai, A., Itakura, S., Yamazaki, M., Tashiro, H., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.		
JOURNAL	NEDO human cDNA sequencing project		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 740)		
TITLE	Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 106-8639, Japan (E-mail:cdna@ime.u-tokyo.ac.jp, Tel:81-3-5449-5586, Fax:81-3-5449-5416)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'-6' and one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).		
FEATURES	Location/Qualifiers		
SOURCE	1. 740		
CDS	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="COLF1505"		
	/tissue_type="colon mucosa"		
	/clone_id="ColF"		
	/note="cloning vector pME18FL3"		
	177..626		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAA91012.1"		
	/db_xref="GI:7020147"		
	/translation="MDALIDLTVKVSPLYLETYGAEITANVLRDMLQEMAGLOLAATTHQGSAAAPGICAPQSAKPGIHFIDIRALIVRVNVEMLDALVGVKVLTDQYQAVRASPTNSKKRKIFSFPPANWTKDILLQALNESQSYLVEDLERS"		
BASE COUNT	161 a 231 c 230 g 118 t		
ORIGIN			
Query Match	96.4%; Score 713; DB 9; Length 740;		
Best Local Similarity	99.7%; Pred. No. 5.7e-103;		
Matches	725; Conservative 0; Mismatches 0; Indels 2; Gaps 1;		
Y	14	AGCGGGGAGACGGGCGGAGAGCGCGGGGATCTTGAGACCTTGGGGCGCGCGCGGAGCG	73
DB	1	AGCGGGGAGACGGGCGGAGAGCGCGGGGATCTTGAGACCTTGGG--GCGGCGCGGAGCG	58
Y	74	CATCTGGATCGCTGAGAACCTGACCGCGAGAGAGCTCAAGAGATTCAAGCTGAAGCT	133
DB	59	CATCTGGATCGCTGAGAACCTGACCGCGAGAGAGCTCAAGAGATTCAAGCTGAAGCT	118
Y	134	GCTGTGGATCGCGCTGCGGAGAGGCTTAAGGGGCGATCCCCGGGGCGCGGCTGTGTGAT	193

Db	119	GCCTGTGGATGACCCCTGGCGGAGGGGCTACAGGGGCCGATCCCGCGGGGCGCGCTGCTGTCCAT	178
QY	194	GAAGCGCTTTGAACCTTCAACGCAAGCTGTGACCTTTACCTTGGAGACTTAACGGCGCCGA	253
Db	179	GGAGCGCTTTGGAGCCTTCAACGCAAGCTGTGATCACTTTACCTTGGAGACTTAACGGCGCCGA	238
QY	254	GCATCAACCGCTTAACGTGTGTGGCGCAACATGGGCTCTGACAGAGATGGACCGGGACGTGCAGGC	313
Db	239	GCTTCAACCGCTTAACGTGTGTGGCGCAACATGGGCTCTGACAGAGATGGACCGGGACGTGCAGGC	298
QY	314	GGCACAGCACAGGGCTCTTGGAGCCGCGCCAGCTGGATCAAGCCCTCTCTCATGTGGC	373
Db	299	GGGCACAGCACAGGGGCTCTTGGAGCCGCGCCAGCTGGATCAAGCCCTCTCTCATGTGGC	358
QY	374	AGCCAAAGCCAGGGCTGTGACTTTATATAGACAGACACCGGGCTGCGCTTATCCGAGAGGTCAC	433
Db	359	AGCCAAAGCCAGGGCTGTGACTTTATATAGACAGACACCGGGCTGCGCTTATCCGAGAGGTCAC	418
QY	434	AAACGTTGAGTGTGCTGTGATGCTCTGTACGGGAAAGTCTTGAACGATAGACATGACA	493
Db	419	AAACGTTGAGTGTGCTGTGATGCTCTGTACGGGAAAGTCTTGAACGATAGACATGACA	478
QY	494	GGCAGTGTGGGGCCGAGGCCCAACCAACCAAGATGCGGAAGTCTTCAAGTTTCAACAC	553
Db	479	GGCAGTGTGGGGCCGAGGCCCAACCAACCAAGATGCGGAAGTCTTCAAGTTTCAACAC	538
QY	554	AGCCTGGAACTGGACCTGCAAGAGACTTGTCTCTACAGAGCCCTTAAGGGAGTCCCAAGTCTTA	613
Db	539	AGCCTGGAACTGGACCTGCAAGAGACTTGTCTCTCTACAGAGCCCTTAAGGGAGTCCCAAGTCTTA	598
QY	614	CTGTGTGAGAGACCTTGGAGCGGAGCTGAGGCTCTCTTCCACGACCAACTCCGGTCAAGCCC	673
Db	599	CTGTGTGAGAGACCTTGGAGCGGAGCTGAGGCTCTCTTCCACGACCAACTCCGGTCAAGCCC	658
QY	674	TGGCAATCCCAACCAATCATCTCGAATCTGATCTTTTATACCAATATACGAAAGCA	733
Db	659	TGGCAATCCCAACCAATCATCTCGAATCTGATCTTTTATACCAATATACGAAAGCA	718
QY	734	GCCTTGAA 740	
Db	719	GCCTTGAA 725	
RESULT 15			
AF310103			
LOCUS	AF310103	712 bp	mRNA linear PRI 04-MAY-2001
DEFINITION	Homo sapiens PYCARD mRNA, complete cds.		
ACCESSION	AF310103		
VERSION	AF310103.1	GI:11096298	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1. (bases 1 to 712)		
JOURNAL	Martinoon, F., Hofmann, K. and Tschopp, J.		
MEDLINE	The pyrin domain: a possible member of the death domain-fold family		
PUBMED	Implicated in apoptosis and inflammation		
REFERENCE	Curr. Biol. 11 (4), R118-R120 (2001)		
AUTHORS	2. (bases 1 to 712)		
TITLE	Martinoon, F., Hofmann, K. and Tschopp, J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (28-SEP-2000) Institute of Biochemistry, University of		
source	Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland		
	Location/Qualifiers		
	1..712		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..588		
	/note="contains PYD and CARD domains"		

```

/codon_start=1
/product="PYCARD"
/protein_id="AAG30286.1"
/db_xref="GI:11096299"

```

KLFSPTDANWTCRDLLOALRESOSYLVEDLERS"

BASE COUNT	166 a	221 c	210 g	115 t
ORIGIN				

Query Match	92.8%;	Score 687;	DB 9;	Length 712;
Best Local Similarity	100.0%;	Pred. No. 7.2e-99;		
Matches 687;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	54	ATGGGCGCGCGCGCAGCCCATCTCGATGCGCTGAGAACTGACCGCGAAGACTC	113
Db	1	ATGGGCGCGCGCGCAGCCCATCTCGATGCGCTGAGAACTGACCGCGAAGACTC	60
QY	114	AGAAATTCAAGCTGAAGCTCTGTTCGGTCCGCTCGCGAAGGCTACCGGCGATCCG	173
Db	61	AGAAATTCAAGCTGAAGCTCTGTTCGGTCCGCTCGCGAAGGCTACCGGCGATCCG	120
QY	174	CGGGGCGCGCTGCTGCATGAGCGCCTTTGACCTCAACGACAGCTGTAGCTTAC	233
Db	121	CGGGGCGCGCTGCTGCATGAGCGCCTTTGACCTCAACGACAGCTGTAGCTTAC	180
QY	234	CTGGAGACTTACGGCGCCGAGCTCAACGCTAACGTCCTGCGCAATAGGCTTGCAAG	293
Db	181	CTGGAGACTTACGGCGCCGAGCTCAACGCTAACGTCCTGCGCAATAGGCTTGCAAG	240
QY	294	ATGGCGGGGCACTGCAGGGCGGCGCAACGACAGGCTTGAGCGCGGCGACGCGGATC	353
Db	241	ATGGCGGGGAGTGTGAGGGCGGCGCAACGACAGGCTTGAGCGCGGCGACGCGGATC	300
QY	354	CAGGCCCTCTCTCAGTCGGCAGCCCAAGCCAGGCTTGACTTTATAGACAGCACCGGCT	413
Db	301	CAGGCCCTCTCTCAGTCGGCAGCCCAAGCCAGGCTTGACTTTATAGACAGCACCGGCT	360
QY	414	GGCCTTATGGCGAGGTCACAAACGTTGATGAGGCTGTGATGCTCTGTACGGGAAGTC	473
Db	361	GGCCTTATGGCGAGGTCACAAACGTTGATGAGGCTGTGATGCTCTGTGAGGGAAGTC	420
QY	474	CTGACGATGAGCAGTACAGGCAAGTCGCGGCGGAGCCACCAACCAAGCAAGATCGG	533
Db	421	CTGACGATGAGCAGTACAGGCAAGTCGCGGCGGAGCCACCAACCAAGCAAGATCGG	480
QY	534	AAGCTTTCAATTCAACCAAGCTTGGAATTGAACTTGACCTTGCAAGACTTCTCAGGCC	593
Db	481	AAGCTTTCAATTCAACCAAGCTTGGAATTGAACTTGACCTTCAAGGCC	540
QY	594	CTTAAGGAGTCCAGTCTCACTGATGAGAGCACTGAGCGGAGGTAGAGTCTCTTCCA	653
Db	541	CTTAAGGAGTCCAGTCTCACTGATGAGAGCACTGAGCGGAGGTAGAGTCTCTTCCA	600
QY	654	GCAACACTCGGTCAGCCCTGGCAATCCACCAATCATCTGATCTGATCTTTTAT	713
Db	601	GCAACACTCGGTCAGCCCTGGCAATCCACCAATCATCTGATCTGATCTTTTAT	660
QY	714	ACAACAATATACAAAAGCCAGCTTGA	740
Db	661	ACAACAATATACAAAAGCCAGCTTGA	687

Search completed: January 29, 2004, 09:40:36
Job time : 2665.52 secs

XX (MILL-) MILLENNIUM PHARM INC.
 PA Bertin J;
 XX
 XX
 XX
 DR WPI: 2001-061973/07.
 DR P-PSDB; AAB20085.
 XX
 PT Isolated intracellular proteins predicted to be involved in regulating
 PT caspase activation are used for diagnosis and treatment of e.g. cancer,
 PT viral infections, autoimmune diseases, neurological diseases and
 PT hematological disorders -
 XX
 XX
 PS Claim 1(a); Fig 21, 208pp; English.
 CC The present sequence is that of cDNA encoding human caspase
 CC recruitment domain 5 (CARD-5, see AAB20085). The cDNA was isolated
 CC from a testis cDNA library using murine CARD-1. Plasmid BPHCS
 CC containing CARD-5 cDNA is deposited as ATCC PTA-213. CARD-5
 CC is an intracellular protein predicted to be involved in regulating
 CC caspase activation. It is useful as a modulating agent in
 CC regulating cellular processes include cell growth and cell death.
 CC Methods of diagnosing and treating patients suffering from a
 CC disorder associated with an abnormal level or rate of apoptotic
 CC cell death, abnormal activity of the Fas/PO-1 receptor complex,
 CC abnormal activity of the tumour necrosis factor receptor complex
 CC or abnormal activity of a caspase involve administering a compound
 CC that modulates the expression or activity of CARD-3, CARD-4, CARD-5
 CC or CARD-6 e.g. a small molecule, antisense nucleic acid, ribozyme
 CC or polypeptide. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, hematological
 CC disorders, inflammatory disorders and immune disorders. CARD
 CC nucleic acids can be used to express CARD proteins in a host cell
 CC e.g. for gene therapy applications, to detect a genetic lesion and
 CC to modulate CARD activity.
 XX
 XX
 SQ Sequence 740 BP, 146 A, 238 C, 236 G, 120 T, 0 other;
 Query Match 100.0%; Score 740; DB 22; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1,7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGCTGCAAGCGGGGTGAGCGGCGGAGCGCGGGAATCTGTGAGCCATGCGGC 60
 DB 1 CGCGTCCGCTGCAAGCGGGGTGAGCGGCGGAGCGCGGGAATCTGTGAGCCATGCGGC 60
 QY 61 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 DB 61 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 TCAAGCTGAAGTGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTG 180
 DB 121 TCAAGCTGAAGTGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTG 180
 QY 181 CGCTGCTGTCTCATGAGCGCTTGTGAGCTTCAACCAAGAGTGTGAGCTTCACTTGAGAG 240
 DB 181 CGCTGCTGTCTCATGAGCGCTTGTGAGCTTCAACCAAGAGTGTGAGCTTCACTTGAGAG 240
 QY 241 CCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 DB 241 CCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 301 GAGGAGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 301 GAGGAGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 QY 361 CTCTCTAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 DB 361 CTCTCTAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 QY 421 TCGCGAGGAGTCAAAAGCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 480
 DB 421 TCGCGAGGAGTCAAAAGCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 480

QY 481 ATGAGCAGTACGAGGAGTGTGCGGCGGAGCCCAACCAAGCAAGATGCGGAGCTCT 540
 DB 481 ATGAGCAGTACGAGGAGTGTGCGGCGGAGCCCAACCAAGCAAGATGCGGAGCTCT 540
 QY 541 TCAAGTTCACACCAAGCTTGAAGTGTGAGCTTGTCTCTCCAGAGCCCTTAAGGG 600
 DB 541 TCAAGTTCACACCAAGCTTGAAGTGTGAGCTTGTCTCTCCAGAGCCCTTAAGGG 600
 QY 601 AGTCCAGGCTTCACTGTGTGAGAGACTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAG 660
 DB 601 AGTCCAGGCTTCACTGTGTGAGAGACTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAG 660
 QY 661 TCGGTCAGAGCCCTGAGCAATCCCAACCAATTCATCTGATCTGATCTTTTATACAAAT 720
 DB 661 TCGGTCAGAGCCCTGAGCAATCCCAACCAATTCATCTGATCTGATCTTTTATACAAAT 720
 QY 721 ATACGAAAAGCCAGCTTGA 740
 DB 721 ATACGAAAAGCCAGCTTGA 740

RESULT 2
 ABR87966
 ID ABR87966 standard; cDNA; 740 BP.
 XX
 AC ABR87966;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 XX Human caspase recruitment domain-5 (CARD-5) cDNA.
 DE
 XX Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; hematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 54..641
 FT /tag= a
 FT /product= "CARD-5"
 PN MO200244354-A2.
 XX
 XX 06-JUN-2002.
 PD
 XX 29-NOV-2001; 2001WC-US44894.
 PF
 XX 01-DEC-2000; 2000US-0728721.
 PR 24-APR-2001; 2001US-0841879.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Bertin J;
 XX
 XX WPI: 2002-557538/59.
 DR P-PSDB; AA0939353.
 XX
 XX Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis -
 XX

PS Claim 7; Fig 3; 100bp; English.

CC The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteine) aspartate-specific proteinases are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CARDs and different
 CC subtypes of CARDs may confer binding specificity. CARD-5 is an
 CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B
 CC (NF-kappaB) transcription factor pathway and binds the CARDs of
 CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
 CC activity and NF-kappaB activation, regulate cell growth and cell death
 CC and be used in gene therapy. The CARD-5 polypeptides are useful for
 CC identifying compounds which bind to them and modulate their activity and
 CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
 CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
 CC can be used to treat immune disorders such as chronic inflammatory
 CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
 CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
 CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
 CC (including tuberculosis and lepromatous leprosy) and in screening and
 CC detection assays. Modulators of CARD-5 activity or expression are also
 CC useful for treating autoimmune disorders, such as systemic lupus
 CC erythematosus and arthritis, cell depletion, neurological disorders,
 CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
 CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
 CC aplastic anaemia, myocardial infarction and stroke. The sequence
 CC presented is the human caspase recruitment domain-5 (CARD-5) cDNA.

CC Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match 100.0%; Score 740; DB 24; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CGCGTCCGGCTGACGCGGGGTGAGCGCGCGCGCGGGGAGCTCTGAGACCAATGGGCG 60
DB 1 CGCGTCCGGCTGACGCGGGGTGAGCGCGCGCGCGGGGAGCTCTGAGACCAATGGGCG 60
QY 61 GCGGCGCGGACGCCATCTGTGATGCGCTGAGAACTTGACCGCGAGAGCTCAAGAAGT 120
DB 61 GCGGCGCGGACGCCATCTGTGATGCGCTGAGAACTTGACCGCGAGAGCTCAAGAAGT 120
QY 121 TCAAGCTGAACTGCTGTGCGTGGCGCTGCGCGAGGGGCTTAAGGGCGCATCCCGGGGCG 180
DB 121 TCAAGCTGAACTGCTGTGCGTGGCGCTGCGCGAGGGGCTTAAGGGCGCATCCCGGGGCG 180
QY 181 CGGCTGCTCATGAGACCGCTTGAACCTCAAGCAAGCTGTGAGCTTCTTAACCTGAGAGA 240
DB 181 CGGCTGCTCATGAGACCGCTTGAACCTCAAGCAAGCTGTGAGCTTCTTAACCTGAGAGA 240
QY 241 CCTACGGGCGGACCTCAACGCTGACGCTGCGCGAGTGGGCTGCGAGAGATGGGCG 300
DB 241 CCTACGGGCGGACCTCAACGCTGACGCTGCGCGAGTGGGCTGCGAGAGATGGGCG 300
QY 301 GCGAGCTGCGAGCGCGCCACGACCAAGGCTTGAAGCCGCGACGCTGGATTCAGGCCC 360
DB 301 GCGAGCTGCGAGCGCGCCACGACCAAGGCTTGAAGCCGCGACGCTGGATTCAGGCCC 360
QY 361 CTGCTGAGTGGGAGCGACGCGAGGCTTGAAGCTTGAAGCAAGCGGCTGCGGCTA 420
DB 361 CTGCTGAGTGGGAGCGACGCGAGGCTTGAAGCTTGAAGCAAGCGGCTGCGGCTA 420
QY 421 TCGGAGGGGTCAAAAGCTTGAAGTGGCTGTGATGCTCTGATGCGAGAGTCTCTGACGG 480
DB 421 TCGGAGGGGTCAAAAGCTTGAAGTGGCTGTGATGCTCTGATGCGAGAGTCTCTGACGG 480
QY 481 ATGAGCACTACCAAGCACTGCGGGCGGACCGACCAAGCAAGCAAGTGGGAGTCTT 540
DB 481 ATGAGCACTACCAAGCACTGCGGGCGGACCGACCAAGCAAGCAAGTGGGAGTCTT 540

```

```

QY 541 TCAAGTTACACACGAGCTGGAAGTGAAGTGAAGAGAGTCTCTCCAGGCCCTAAGG 600
DB 541 TCAAGTTACACACGAGCTGGAAGTGAAGTGAAGAGAGTCTCTCCAGGCCCTAAGG 600
QY 601 AGTCCAGTCTTACTGTGTGAGAGACCTGAGCGGAGCTAGAGCTCTTCCAGCAAC-660
DB 601 AGTCCAGTCTTACTGTGTGAGAGACCTGAGCGGAGCTAGAGCTCTTCCAGCAAC-660
QY 661 TCCGCTAGGCGCCCTGGGAATCCCAAGCAATCCTGAATCTGATCTTTATACAAAT 720
DB 661 TCCGCTAGGCGCCCTGGGAATCCCAAGCAATCCTGAATCTGATCTTTATACAAAT 720
QY 721 ATAGAAAGCCAGCTTGA 740
DB 721 ATAGAAAGCCAGCTTGA 740

```

RESULT 3
 ID ABR87967/c
 ABR87967 standard; DNA; 740 BP.

AC ABR87967;

DT 07-OCT-2002 (first entry)

XX Human caspase recruitment domain-5 (CARD-5) cDNA complementary strand.

KM Human, ss; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KM immunosuppressive; caspase; cysteine1 aspartate-specific proteinase;
 KM apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KM cell proliferation; gene therapy; immune disorder;

KM chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KM sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KM human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KM lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KM arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KM Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KM myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KM stroke.

OS Homo sapiens.

FN WO200244354-A2.

PD 06-JUN-2002.

XX 29-NOV-2001; 2001WO-US44894.

PF 01-DEC-2000; 2000US-0728721.

PR 24-APR-2001; 2001US-0841879.

XX (MILL-) MILLENIUM PHARM INC.

PA Bertin J;

PI WPI; 2002-557538/59.

XX Novel isolated murine or human caspase recruitment domain (CARD)-5

PT polypeptide, useful for treating immune disorders such as Hashimoto's

PT thyroiditis, graft rejection, allergy, glomerular nephritis,

PT tuberculosis

PS Claim 8; Fig 3; 100bp; English.

CC The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteine) aspartate-specific proteinases are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CARDs and different
 CC subtypes of CARDs may confer binding specificity. CARD-5 is an
 CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B

CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC encoding human CARD-5 is used to identify cDNA encoding CARD-7.

XX Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match 100.0%; Score 740; DB 25; Length 740;

Best Local Similarity 100.0%; Pred. No. 1.7e-147;

Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTCCGCTGACGCGGGGTGAGCGGCGGCGGCGGATCTTGAGCCATGGGCG 60
 DB 1 CGGCTCCGCTGACGCGGGGTGAGCGGCGGCGGCGGATCTTGAGCCATGGGCG 60
 QY 61 GCGGCGCGGCGGCGGATCTTGAGCGGCGGCGGCGGATCTTGAGCGGCGGCG 120
 DB 61 GCGGCGCGGCGGCGGATCTTGAGCGGCGGCGGCGGATCTTGAGCGGCGGCG 120
 QY 121 TCAAGCTGAAGCTGCTGCGGTGCGGCTGCGGCGGCGGCTACGGGCGGCGGCG 180
 DB 121 TCAAGCTGAAGCTGCTGCGGTGCGGCTGCGGCGGCGGCTACGGGCGGCGGCG 180
 QY 181 CGCTGCTGTCATGAGCGCTTGAGCTCAACGCAAGCTGCTGCTTCACTTGAGA 240
 DB 181 CGCTGCTGTCATGAGCGCTTGAGCTCAACGCAAGCTGCTGCTTCACTTGAGA 240
 QY 241 CTTACGCGCGCGGCGGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 300
 DB 241 CTTACGCGCGCGGCGGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 300
 QY 301 GGCAGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 301 GGCAGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 QY 361 CTCTCTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420
 DB 361 CTCTCTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420
 QY 421 TCGGAGAGGCTCAAAAGCTGAGTGGCTGCTGCTTATAGGAGAGGCTCTGACG 480
 DB 421 TCGGAGAGGCTCAAAAGCTGAGTGGCTGCTGCTTATAGGAGAGGCTCTGACG 480
 QY 481 ATGAGAGGTCACGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540
 DB 481 ATGAGAGGTCACGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540
 QY 541 TCAAGTTTCAACGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 600
 DB 541 TCAAGTTTCAACGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 600
 QY 601 AGTCCGAGCTTACTGCTGAGAGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAG 660
 DB 601 AGTCCGAGCTTACTGCTGAGAGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAG 660
 QY 661 TCCGCTGAGCGGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAG 720
 DB 661 TCCGCTGAGCGGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAG 720

DB 661 TCCGCTGAGCGGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 720
 QY 721 ATACGAAAAGCCGAGCTTGAA 740
 DB 721 ATACGAAAAGCCGAGCTTGAA 740

RESULT 5

AAD03890 standard; cDNA; 770 BP.

AC AAD03890;

DT 02-JUL-2001 (first entry)

DE Human target of methylation-induced silencing-1 (TMS1) cDNA.

KM Human target of methylation-induced silencing-1; TMS1; cytosolic;

KM antiproliferative; apoptosis inducer; gene therapy; CpG island;

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS /tag= a /product= "Human TMS1 protein"

PN WO200129235-A2.

PD 26-APR-2001.

PF 18-OCT-2000; 2000WO-US28747.

PR 18-OCT-1999; 99US-0159975.

PA (YEM-) UNIV EMORY.

PI Vertino PM;

DR WPI; 2001-290922/30.

DR P-PSDB; AAB00588.

PT Novel gene TMS1, transcriptionally silenced due to increased

PT methylation useful for identifying subject at risk of developing tumor

PT characterized by abnormal methylation, for treating cancer by inducing

PS Claim 68; Page 113-114; 124pp; English.

CC The invention relates to identification of target of methylation-induced

CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to

CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1

CC consists of a carboxy terminal caspase-recruiting domain (CARD) and

CC plays a role in induction of apoptosis. TMS1 gene and protein are useful

CC as tools for diagnosing and treating a subject at risk of developing

CC cancer (e.g. breast cancer) characterized by abnormal CpG methylation or

CC abnormally low levels of TMS1 expression products. Unique fragments of

CC TMS1 molecule are used as probes. TMS1 gene is useful in gene therapy.

CC TMS1 molecule is also useful for treating abnormal cell proliferation by

CC increasing TMS1 polypeptide level to an above normal level. The CpG

CC island region of TMS1 or its fragments are used to study the methylation

CC patterns apart from any coding region contained in it.

CC The present sequence is human target of methylation-induced silencing-1

CC (TMS1) cDNA.

SO Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

Query Match 99.2%; Score 734; DB 22; Length 770;

Best Local Similarity 100.0%; Pred. No. 3.2e-146;

Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGACGCGGAGTGAAGCGGCGGCGGCGGCGGATCTTGAGAGCATGGGCGGCGG 66

Db 28 CGGCTGACAGGGGGTGAAGCCGCGGACGCGGATCTTGAAGCCATGAGGCGCGCC
 Qy 67 GCGAGCCCATCTTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 126
 Db 88 GCGAGCCCATCTTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 147
 Qy 127 TGAAGCTGCTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 186
 Db 148 TGAAGCTGCTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 207
 Qy 187 TGTTCATGAGAGCTTGAGAACTGACCGAGAGCTGAGCTTCACTTGAAGCTTACG 246
 Db 208 TGTTCATGAGAGCTTGAGAACTGACCGAGAGCTGAGCTTCACTTGAAGCTTACG 267
 Qy 247 GCGCCGAGCTACCGCTTACGCTGCGGAGATGAGGCTTGAAGAGATGCGGAGAGC 306
 Db 268 GCGCCGAGCTACCGCTTACGCTGCGGAGATGAGGCTTGAAGAGATGCGGAGAGC 327
 Qy 307 TGAAGCGGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 366
 Db 328 TGAAGCGGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 387
 Qy 367 AGTCGAGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 426
 Db 388 AGTCGAGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 447
 Qy 427 GGGTCAAAAGCTTGAAGTGTGCTGAGTCTTGAAGAGAGTCTTGAAGAGATGAGC 486
 Db 448 GGGTCAAAAGCTTGAAGTGTGCTGAGTCTTGAAGAGAGTCTTGAAGAGATGAGC 507
 Qy 487 AGTACCAAGAGAGCGGCGGAGAGCTTGAAGAGAGAGTCTTGAAGAGATGAGC 546
 Db 508 AGTACCAAGAGAGCGGCGGAGAGCTTGAAGAGAGAGTCTTGAAGAGATGAGC 567
 Qy 547 TCACACGAGCGTGAAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 606
 Db 568 TCACACGAGCGTGAAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 627
 Qy 607 AGTCTCACTGCTGAGAGAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 666
 Db 628 AGTCTCACTGCTGAGAGAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 687
 Qy 667 CAGCCGCGGAGAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 726
 Db 688 CAGCCGCGGAGAGCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAGAGAGTCTTGAAG 747
 Qy 727 AAAGCAGCTTGAAG 740
 Db 748 AAAGCAGCTTGAAG 761
 RESULT 6
 AAL47126
 ID AAL47126 standard; DNA: 782 BP.
 XX AAL47126;
 AC
 XX 20-AUG-2002 (first entry)
 DT
 XX
 DE Pycin domain containing protein Pycard coding sequence.
 XX
 XX Pycin domain; PYD domain; antiinflammatory; antiParkinsonian;
 KM antiarteriosclerotic; antiapoptotic; antibacterial; virocidic;
 KM neuroprotective; antiallergic; antirheumatic; antiasthmatic;
 KM nephroprotective; osteoprotective; intracellular signal transduction;
 KM inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KM arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KM osteoarthritis; glomerulonephritis; gene; ds.
 KM
 XX Unidentified.
 OS
 XX
 PN MO200240668-A2.

XX
 PD 23-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-EP12545.
 XX
 PR 15-NOV-2000; 2000DE-1056687.
 PR 30-NOV-2000; 2000DE-1059595.
 XX
 PA (APOT-) APOTEC RES & DEV LTD.
 XX
 PI Techopp J, Martillon F;
 XX
 DR WPI: 2002-427093/45.
 DR P-PSDB; AA017854.
 XX
 PT New DNA encoding protein with pycin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies -
 XX
 PS Claim 5, Fig 1, 116p; German.
 XX
 CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention.
 XX
 SQ Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other;
 Query Match 99.2%; Score 734; DB 24; Length 782;
 Best Local Similarity 100.0%; Pred. No. 3.2e-146;
 Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 CCGCTGACAGGGGGTGAAGCCGCGGACGCGGATCTTGAAGCCATGAGGCGCGCC 66
 Db 40 CCGCTGACAGGGGGTGAAGCCGCGGACGCGGATCTTGAAGCCATGAGGCGCGCC 99
 Qy 67 GCGAGCCCATCTTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 126
 Db 100 GCGAGCCCATCTTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 159
 Qy 127 TGAAGCTGCTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 186
 Db 160 TGAAGCTGCTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGATTCAAGC 219
 Qy 187 TGTTCATGAGAGCTTGAGAACTGACCGAGAGCTGAGCTTCACTTGAAGCTTACG 246
 Db 220 TGTTCATGAGAGCTTGAGAACTGACCGAGAGCTGAGCTTCACTTGAAGCTTACG 279
 Qy 247 GCGCCGAGCTACCGCTTACGCTGCGGAGATGAGGCTTGAAGAGATGCGGAGAGC 306
 Db 280 GCGCCGAGCTACCGCTTACGCTGCGGAGATGAGGCTTGAAGAGATGCGGAGAGC 339
 Qy 307 TGAAGCGGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 366
 Db 340 TGAAGCGGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 399
 Qy 367 AGTCGAGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 426
 Db 400 AGTCGAGCGGAGAGCTTGAGAACTGACCGCGGAGAGCTTGAAGAGATGCGGAGAGC 459
 Qy 427 GGGTCAAAAGCTTGAAGTGTGCTGAGTCTTGAAGAGAGTCTTGAAGAGATGAGC 486
 Db 460 GGGTCAAAAGCTTGAAGTGTGCTGAGTCTTGAAGAGAGTCTTGAAGAGATGAGC 519
 Qy 487 AGTACCAAGAGAGCGGCGGAGAGCTTGAAGAGAGAGTCTTGAAGAGATGAGC 546
 Db 520 AGTACCAAGAGAGCGGCGGAGAGCTTGAAGAGAGAGTCTTGAAGAGATGAGC 579

QY 608 GTCTACCTGTGAGAGACCTGAGCGAGCTGAGGCTCTCCACAGCAACACTCCGGTC 667
 DB 602 GTCTACCTGTGAGAGACCTGAGCGAGCTGAGGCTCTCCACAGCAACACTCCGGTC 661
 QY 668 AGCCCTGAGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 727
 DB 662 AGCCCTGAGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 721
 QY 728 AAGCCAGCTTGAA 740
 DB 722 AAGCCAGCTTGAA 734

RESULT 8

AAZ33631
 ID AAZ33631 standard; cDNA; 779 BP.

AC AAZ33631;

DT 08-DEC-1999 (first entry)

DE Human breast tumour-associated EST 21.

XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;
 KM treatment; tumour; cytostatic; medicament; ss.

XX Homo sapiens.

PN DE19813839-A1.

PD 23-SEP-1999.

PF 20-MAR-1998; 98DE-1013839.

PR 20-MAR-1998; 98DE-1013839.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI; 1999-528981/45.

PT Human nucleic acid sequences and protein products from tumor breast
 tissue, useful for breast cancer therapy -

PS Claim 1a; 101; 188pp; German.

XX This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC actively against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AAZ33631-248617 represents expressed
 CC sequence tags described in the method of the invention.

XX Sequence 779 BP; 149 A; 252 C; 251 G; 127 T; 0 other;

QY Query Match 99.1%; Score 733; DB 20; Length 779;

DB Best Local Similarity 100.0%; Pred. No. 5.3e-146;
 Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGAGCGGGGTGAGCGCGCGGACCGGGAATCTTGAGCCATGGGGCGCGCG 66
 DB 47 CGGCTGAGCGGGGTGAGCGCGCGGACCGGGAATCTTGAGCCATGGGGCGCGCG 106
 QY 67 GCGACGCGATCTGAGATGCGCTGAGAACTGACCGCGCGAGAGCTCAAGATTCAAGC 126
 DB 107 GCGACGCGATCTGAGATGCGCTGAGAACTGACCGCGCGAGAGCTCAAGATTCAAGC 166
 QY 127 TGAAGCTGCTGTGCGTCCGCTGCGGAGAGGCTACGCGCGCGATCCCGGGCGCGCTGC 186
 DB 167 TGAAGCTGCTGTGCGTCCGCTGCGGAGAGGCTACGCGCGCGATCCCGGGCGCGCTGC 226

QY 187 TGTTCATGAGACGCTTGGACCTTCACCGACAGAGTGTGACTTCTACCTTGAGAACTTACG 246
 DB 227 TGTTCATGAGACGCTTGGACCTTCACCGACAGAGTGTGACTTCTACCTTGAGAACTTACG 286
 QY 247 GCGCCGAGCTCACCGCTTAAGTGTCTGCGGACATGAGGCTGAGAGATGAGCGCGGACG 306
 DB 287 GCGCCGAGCTCACCGCTTAAGTGTCTGCGGACATGAGGCTGAGAGATGAGCGCGGACG 346
 QY 307 TGCAGGCGGCGACGACCAAGGCTCTGAGACCGCGCTGAGTTCAGAGCCCTCTCTC 366
 DB 347 TGCAGGCGGCGACGACCAAGGCTCTGAGACCGCGCTGAGTTCAGAGCCCTCTCTC 406
 QY 367 AGTCGGCAGCGCAAGCCAGGCTTGCATTTATACCAAGCAACCGGCTTGGCTTATGGCA 426
 DB 407 AGTCGGCAGCGCAAGCCAGGCTTGCATTTATACCAAGCAACCGGCTTGGCTTATGGCA 466
 QY 427 GGGTCAAAACGTTGAGTGTGCTGATGATGCTCTGTACGGGAAAGTCTTGACGATGAGC 486
 DB 467 GGGTCAAAACGTTGAGTGTGCTGATGATGCTCTGTACGGGAAAGTCTTGACGATGAGC 526
 QY 487 AGTACGAGCAGTGCAGGCTCGAGCCGACCAACCAAGCAAGATGCGGAGCTTTCAGTT 546
 DB 527 AGTACGAGCAGTGCAGGCTCGAGCCGACCAACCAAGCAAGATGCGGAGCTTTCAGTT 586
 QY 547 TCACACCAAGCTTGAAGTGTGACCTGCAAGAGCTTCTCTCCAGGCTTAAAGGAGTCCC 606
 DB 587 TCACACCAAGCTTGAAGTGTGACCTGCAAGAGCTTCTCTCCAGGCTTAAAGGAGTCCC 646
 QY 607 AGTCTACCTGTGAGAGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
 DB 647 AGTCTACCTGTGAGAGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 706
 QY 667 CAGCCCTTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACCAATATACGA 726
 DB 707 CAGCCCTTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACCAATATACGA 766
 QY 727 AAGCCAGCTTGA 739
 DB 767 AAGCCAGCTTGA 779

RESULT 9

AAH34052
 ID AAH34052 standard; cDNA; 806 BP.

AC AAH34052;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:1134.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma; ss.

XX Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000MO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.

XX P-PSDB; AAG74647.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PS Claim 1; Page 2955; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG7788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 806 BP; 163 A; 253 C; 260 G; 128 T; 2 other;
 Query Match 97.5%; Score 721.2; DB 22; Length 806;
 Best Local Similarity 99.6%; Pred. No. 1.7e-143;
 Matches 731; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 7 CGGCTGAGGCGGGGTGAGCGCGCGGAGCGCGGGGAGTCTTGAGAGCAATGGCGGCGCGC 66
 68 CGGCTGAGGCGGGGTGAGCGCGCGGAGCGCGGGGAGTCTTGAGAGCAATGGCGGCGCGC 127
 67 GCGAGCGCATCTCTGATGTCGCTGAGAACTGACCGCGCGAGAGCTCAAGAGTTCAAGC 126
 128 GCGAGCGCATCTCTGATGTCGCTGAGAACTGACCGCGCGAGAGCTCAAGAGTTCAAGC 187
 127 TGAAGCTGCTGTGCGTGCCTGCGCGAGGAGCTACGCGGCGGCGCGCGCTGC 186
 188 TGAAGCTGCTGTGCGTGCCTGCGCGAGGAGCTACGCGGCGGCGCGCGCTGC 247
 187 TGTCCATGAGAGCGCTTGAGACTCAACGAGCTGTGATGCTTACTTGAGAGCTAAGC 246
 248 TGTCCATGAGAGCGCTTGAGACTCAACGAGCTGTGATGCTTACTTGAGAGCTAAGC 307
 247 GCGCGGAGCTCAACGCTGTCGCGCAGATGAGGCTGACAGAGATGAGCGCGGAGC 306
 308 GCGCGGAGCTCAACGCTGTCGCGCAGATGAGGCTGACAGAGATGAGCGCGGAGC 367
 307 TGCAGGCGGCGCAGCAGCAGGAGCTCTGAGAGCGCGCGCAGCTGAGATCCAGGCGCTCTC 366
 368 TGCAGGCGGCGCAGCAGCAGGAGCTCTGAGAGCGCGCGCAGCTGAGATCCAGGCGCTCTC 427
 367 AGTGGGAGCGCAGCAGCAGGAGCTCTGAGAGCGCGCGCAGCTGAGATCCAGGCGCTCTC 426
 428 AGTGGGAGCGCAGCAGCAGGAGCTCTGAGAGCGCGCGCAGCTGAGATCCAGGCGCTCTC 487
 427 GGGTCACAAAGCTTGAAGTGTGCTGAGAGCTCTGTAACGAGAGAGTCTCTGAGATGAGC 486
 488 GGGTCACAAAGCTTGAAGTGTGCTGAGAGCTCTGTAACGAGAGAGTCTCTGAGATGAGC 547
 487 AGTACCAAGCAGTGGGCGGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTTCAGTT 546
 548 AGTACCAAGCAGTGGGCGGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTTCAGTT 607
 547 TCAACACGAGCTGGAATGAGCTGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTTCAGTT 606
 608 TCAACACGAGCTGGAATGAGCTGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTTCAGTT 667
 607 AGTCTTACTGTGTGAGAGCCTGAGAGCGAGAGCTGAGGCTCTCTTCCAGCAACACTCCGGT 666

Db 668 AGTCTTACTGTGTGAGAGCCTGAGAGCGAGCTGAGGCTCTCTCCAGCAACACTCCGGT 727
 Qy 667 CAGCCCTGCGCAATCCCAATCAATCTGATCTTTTATATACATATATCGA 726
 Db 728 CA-CCCCCTGGCAATCCCAATCAATCTGATCTTTTATATACATATATCGA 786
 Qy 727 AAAGCCAGCTTGA 740
 Db 787 AAAGCCAGCTTGA 800
 RESULT 10
 AACT7884
 ID AACT7884 standard; cDNA; 811 BP.
 XX
 AC AACT7884;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:278.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
 KW vasoconstrictor; antiproliferative; angiogenic; gene therapy; inflammation;
 KW immune reaction; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 OS Homo sapiens.
 XX
 XX MO20005350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000MO-US05882.
 XX
 XX 12-MAR-1999; 99US-0124270.
 PR
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-58753/55.
 DR P-PSDB; AAB3675.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 PS Claim 1; Page 841; 2352pp; English.
 PS
 CC AACT7607 to AACT8448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC nocotropic; vasoconstrictor; antiproliferative; angiogenic; gene therapy;
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AACT8449 to

AC AAC98638;
 XX
 DT 09-MAR-2001 (first entry)
 XX Human colon cancer antigen nucleotide sequence SEQ ID NO:648.
 XX
 DE Human colon cancer antigen nucleotide sequence SEQ ID NO:648.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW identification; cytoskeletal; cardioactive; neuroprotective; vulnary;
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;
 KW nephrotropic; anti-infective; antibacterial; gene therapy; wound;
 KW neural disorder; immune system disorder; muscular disorder;
 KW reproductive disorder; gastrointestinal disorder; renal disorder;
 KW infectious disease; cardiovascular disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200055351-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05883.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-587534/55.
 XX
 DR P-PSDB; AAB53881.
 XX
 XX
 PT Colon cancer associated gene sequences, referred to as colon cancer
 PT antigen, useful for the treatment, prevention, and diagnosis of colon
 PT disorders such as colon cancer -
 XX
 PS Claim 1; Page 1175; 2104pp; English.
 XX
 CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
 CC human colon cancer antigens can have cytoskeletal, cardioactive, muscular;
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
 CC vulnary, nephrotropic, anti-infective and antibacterial activities, and
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,
 CC proteins and antibodies to the proteins are useful for the prevention,
 CC treatment and diagnosis of colon disorders, such as colon cancer. The
 CC polynucleotides may be used in diagnostics, such as for colon cancer.
 CC chromosome identification, and as hybridisation probes. The proteins
 CC may also be used to prevent diseases such as neural disorders, immune
 CC system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, wounds, renal disorders, infectious
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
 CC AAB54007 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 432 BF; 77 A; 142 C; 152 G; 61 T; 0 other;
 XX
 Query Match 52.0%; Score 384.6; DB 21; Length 432;
 Best Local Similarity 98.4%; Pred. No. 2.7e-72;
 Matches 420; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
 QY 9 GGTGACGCGGGGAGCGGCGGCGGCGGCGGAGTCTGAGACCAATGGGGCGCGGGCGG 68
 DB 8 GGTGACGCGGGGAGCGGCGGCGGCGGCGGAGTCTGAGACCAATGGGGCGCGGGCGG 67
 QY 69 GACGCCATCTGAGTGCCTGAGAACTGACCGCGGAGAGCTCAAGAAATTCAGACTG 128
 DB 68 GACGCCATCTGAGTGCCTGAGAACTGACCGCGGAGAGCTCAAGAAATTCAGACTG 127
 QY 129 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188
 DB 128 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
 QY 189 TCATGAGAGCGCTTGAAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248

DB 188 TCATGAGAGCGCTTGAAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247
 QY 249 GCCAGCTCACCCTTAACCTGCTGCGCGACATGAGGCTTCAGAGATGCGCGGACGCTG 308
 DB 248 GCCAGCTCACCCTTAACCTGCTGCGCGACATGAGGCTTCAGAGATGCGCGGACGCTG 307
 QY 309 CAGGCGGCGCACCGACAGGAGCTCTGAGACCGCGCGCACTGGATTCAGAGCCCTCTCAG 368
 DB 308 CAGGCGGCGCACCGACAGGAGCTCTGAGACCGCGCGCACTGGATTCAGAGCCCTCTCAG 366
 QY 369 TCGGACGACCAAGCGAGCTGCACTTT-ATAGACAGACCGGAGCTGCGCTTATGCGGAG 427
 DB 367 TCGGACGACCAAGCGAGCTGCACTTTATAGACAGACCGGAGCTGCGCTTATGCGGAG 425
 QY 428 GGTCAACA 434
 DB 426 GGTCAACA 432
 RESULT 14
 AAD03907
 ID AAD03907 standard; cDNA; 405 BP.
 XX
 AC AAD03907;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE Alternatively spliced form of human TMS1 cDNA (lacking exon3).
 XX
 KW Human; target of methylation-induced silencing-1; TMS1; cytoskeletal;
 KW anti-proliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 75..404
 FT /tag= a
 FT /product= "Alternatively spliced form of human TMS1
 FT protein lacking exon3"
 FT /note= "CDS does not include stop codon"
 FT /partial
 XX
 PN WO200129235-A2.
 XX
 PD 26-APR-2001.
 XX
 PF 18-OCT-2000; 2000WO-US28747.
 XX
 PR 18-OCT-1999; 99US-0159975.
 XX
 PA (UYEW-) UNIV EMORY.
 XX
 PI Vertino PM;
 XX
 DR WPI: 2001-290922/30.
 XX
 DR P-PSDB; AAE00595.
 XX
 PT Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis -
 XX
 PS Claim 106; Page 123-124; 124pp; English.
 XX
 CC The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or

CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is alternatively spliced form of human target of
CC methylation-induced silencing-1 (TMS1) cDNA lacking exons.
CC Note: The present sequence designated as SEQ ID NO:26, is referred
CC as DNA throughout the specification, however this sequence has been
CC referred as protein in claim 106.

SQ Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other;

Query Match	51.1%;	Score 378;	DB 22;	Length 405;
-------------	--------	------------	--------	-------------

Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	CGGCTGACGCGGGGTGAGCGCGGACGCGGCCGCGGGATCTCGAGCCATGAGGGCGCGGC	66
Db	CGGCTGACGCGGGGTGAGCGCGGCGGCGCGGGATCTCGAGCCATGAGGGCGCGGC	28
QY	GCGACGCCATCTTGAGATGCGCTGGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGC	126
Db	GCGACGCCATCTTGAGATGCGCTGGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGC	88
QY	TGAAGCTGCTGCGGTGCGGTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	186
Db	TGAAGCTGCTGCTGCGGTGCGGTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	148
QY	TGTCATGAGACGCGCTTGGACCTCAACGCAAGCTGTGAGCTTCTACCTGAGACCTTACG	246
Db	TGTCATGAGACGCGCTTGGACCTCAACGCAAGCTGTGAGCTTCTACCTGAGACCTTACG	208
QY	GCGCGGAGCTCACCGCTAACGTGCTGCGCGACATGAGGCTTGAGGAGATGCGCGGGCAGC	306
Db	GCGCGGAGCTCACCGCTAACGTGCTGCGCGACATGAGGCTTGAGGAGATGCGCGGGCAGC	268
QY	TGCAGGGGGCCACGACCAAGGGCTTGGAGCCGCGCCAGCTTGGGATCCAGAGCCCTCTCTC	366
Db	TGCAGGGGGCCACGACCAAGGGCTTGGAGCCGCGCCAGCTTGGGATCCAGAGCCCTCTCTC	328
QY	AGTGGAGACCCAGACCC	384
Db	AGTGGAGACCCAGACCC	388

RESULT 15

ID	standard; DNA; 2821 BP.
ADD03889	

AC	AAD03889;
XX	
DT	02-JUL-2001 (first entry)
DT	

Human target of methylation-induced silencing-1 (TMS1) genomic DNA.

Human; target of methylation-induced silencing-1; TMS1; cytostatic;

KW caspase-recruiting domain; CARD; cancer; breast; ds.

OS Homo sapiens.

FH	Key	Location/Qualifiers
		100-170-

1
/*lag= a
/lag=

FT	1256	3261	/note= "This region is specifically claimed in claim 71"
CD	1256	3261	

```
FT      / *tag= b  
EM      /endjust = "ifrom mem section"
```

FT	/note="CD
FT	1356 1539

FT	/+tag= c
FT	/number= 1
FT	1530..1742
FT	/+tag= d
FT	/number= 1
FT	1743..1799
FT	/+tag= e
FT	/number= 2
FT	1800..2104
FT	/+tag= f
FT	/number= 2
FT	2105..2361
FT	/+tag= g
FT	/number= 3

PN W0200129235-A2.

PD 26-APR-2001.

PF 18-OCT-2000; 2000WO-US28747.

PR 18-OCT-1999; 99US-0159975.

PA (UYEM-) UNIV EMORY.

PI Vertino PM;

WPI; 2001-290922/30.

XX

PT methylation useful for identifying subject at risk of develop

PT apoptosis -

PS Claim 68; Page 112-113; 124pp; English.

The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression, products, unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy. TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it. The present sequence is human target of methylation-induced silencing-1 (TMS1) genomic DNA.

SQ Sequence 2821 BP; 614 A; 864 C; 757 G; 586 T; 0 other;

Query Match 48.6%; Score 360; DB 22; Length 2821;

Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	Db
381	2101
CCAGGCGCTGCACTTTATAGACGACGACCGCGGCTGCGCTTATCGGAGGGTCACAAACGTT	CCAGGCGCTGCACTTTATAGACGACGACCGCGGCTGCGCTTATCGGAGGGTCACAAACGTT
440	2166

Oy 441 GAGTGGCTGCTGATGCTCTGACGGGAGAAGTCTGACGGATGACAGTACCAGGCACTG 500
 Db 2161 GAGTGGCTGCTGATGCTCTGACGGGAGAAGTCTGACGGATGACAGTACCAGGCACTG 2222

501 CGGGCCGAGCCCAACCCAGCAAGATGCGGAAGCTTTCAGTTTCACCCAGCCTGG 560
2221 CGGGCCGAGCCCAACCCAGCAAGATGCGGAAGCTTTCAGTTTCACCCAGCCTGG 2280

561 AACTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAGGGAGTCCAGTCTACCTGGTG 620

Db	2281	AACTGACCTGCAAGACTTCTCTCCAGGCCCTAAGGAGTCCAGTCTTACTGTG	2340
Qy	621	GAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACA	680
Db	2341	GAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACA	2400
Qy	681	CCGACCAATCATCTGATCTTTTATACACATATACG	740
Db	2401	CCGACCAATCATCTGATCTTTTATACACATATACG	2460

Search completed: January 29, 2004, 03:29:23
 Job time : 184.487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model 1

Run on: January 29, 2004, 03:01:52 (without time 1474.73 Seconds
12195.625 Million cell updates/sec)

Title: US-09-996-617-7

Perfect score: 740
Sequence: 1 cgcgtccggtcgcagcg995.....atacgaagcagcttga 740

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estm:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vtc:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729.4	98.6	930	12	BM459678 AGENCOURT
2	727	98.2	804	12	BM926438 AGENCOURT
3	722	97.6	767	12	BM910918 AGENCOURT
4	721.2	97.5	792	12	BM549665 AGENCOURT

5	713.4	96.4	750	12	BM764161
6	704.8 <td>95.2 <td>725 <td>14 <td>CA309883</td> </td></td></td>	95.2 <td>725 <td>14 <td>CA309883</td> </td></td>	725 <td>14 <td>CA309883</td> </td>	14 <td>CA309883</td>	CA309883
7	690.8 <td>93.4 <td>836 <td>12 <td>BM767422</td> </td></td></td>	93.4 <td>836 <td>12 <td>BM767422</td> </td></td>	836 <td>12 <td>BM767422</td> </td>	12 <td>BM767422</td>	BM767422
8	688.6 <td>93.1 <td>748 <td>10 <td>BM337806</td> </td></td></td>	93.1 <td>748 <td>10 <td>BM337806</td> </td></td>	748 <td>10 <td>BM337806</td> </td>	10 <td>BM337806</td>	BM337806
9	670.4 <td>90.6 <td>678 <td>9 <td>BM148558</td> </td></td></td>	90.6 <td>678 <td>9 <td>BM148558</td> </td></td>	678 <td>9 <td>BM148558</td> </td>	9 <td>BM148558</td>	BM148558
10	667.6 <td>90.2 <td>741</td> <td>12 <td>BM051141</td> </td></td>	90.2 <td>741</td> <td>12 <td>BM051141</td> </td>	741	12 <td>BM051141</td>	BM051141
11	660 <td>89.2 <td>676 <td>12 <td>BM972785</td> </td></td></td>	89.2 <td>676 <td>12 <td>BM972785</td> </td></td>	676 <td>12 <td>BM972785</td> </td>	12 <td>BM972785</td>	BM972785
12	654.2 <td>88.4 <td>675 <td>10 <td>BM988204</td> </td></td></td>	88.4 <td>675 <td>10 <td>BM988204</td> </td></td>	675 <td>10 <td>BM988204</td> </td>	10 <td>BM988204</td>	BM988204
13	647.4 <td>87.5 <td>744 <td>10 <td>BM255521</td> </td></td></td>	87.5 <td>744 <td>10 <td>BM255521</td> </td></td>	744 <td>10 <td>BM255521</td> </td>	10 <td>BM255521</td>	BM255521
14	639 <td>86.4 <td>687 <td>10 <td>BM909218</td> </td></td></td>	86.4 <td>687 <td>10 <td>BM909218</td> </td></td>	687 <td>10 <td>BM909218</td> </td>	10 <td>BM909218</td>	BM909218
15	617.2 <td>83.4 <td>638 <td>12 <td>BM998354</td> </td></td></td>	83.4 <td>638 <td>12 <td>BM998354</td> </td></td>	638 <td>12 <td>BM998354</td> </td>	12 <td>BM998354</td>	BM998354
16	610 <td>82.4 <td>1079 <td>12 <td>BM456838</td> </td></td></td>	82.4 <td>1079 <td>12 <td>BM456838</td> </td></td>	1079 <td>12 <td>BM456838</td> </td>	12 <td>BM456838</td>	BM456838
17	606 <td>81.9 <td>633 <td>12 <td>BM681310</td> </td></td></td>	81.9 <td>633 <td>12 <td>BM681310</td> </td></td>	633 <td>12 <td>BM681310</td> </td>	12 <td>BM681310</td>	BM681310
18	604.6 <td>81.7 <td>1201 <td>13 <td>BM356773</td> </td></td></td>	81.7 <td>1201 <td>13 <td>BM356773</td> </td></td>	1201 <td>13 <td>BM356773</td> </td>	13 <td>BM356773</td>	BM356773
19	581.2 <td>78.5 <td>830 <td>10 <td>BM238010</td> </td></td></td>	78.5 <td>830 <td>10 <td>BM238010</td> </td></td>	830 <td>10 <td>BM238010</td> </td>	10 <td>BM238010</td>	BM238010
20	553.8 <td>74.8 <td>621 <td>12 <td>BM146829</td> </td></td></td>	74.8 <td>621 <td>12 <td>BM146829</td> </td></td>	621 <td>12 <td>BM146829</td> </td>	12 <td>BM146829</td>	BM146829
21	546 <td>73.8 <td>866 <td>10 <td>BM684386</td> </td></td></td>	73.8 <td>866 <td>10 <td>BM684386</td> </td></td>	866 <td>10 <td>BM684386</td> </td>	10 <td>BM684386</td>	BM684386
22	544.6 <td>73.6 <td>634 <td>9 <td>BM991236</td> </td></td></td>	73.6 <td>634 <td>9 <td>BM991236</td> </td></td>	634 <td>9 <td>BM991236</td> </td>	9 <td>BM991236</td>	BM991236
23	539 <td>72.8 <td>878 <td>13 <td>BM501178</td> </td></td></td>	72.8 <td>878 <td>13 <td>BM501178</td> </td></td>	878 <td>13 <td>BM501178</td> </td>	13 <td>BM501178</td>	BM501178
24	536 <td>72.4 <td>888 <td>10 <td>BM906296</td> </td></td></td>	72.4 <td>888 <td>10 <td>BM906296</td> </td></td>	888 <td>10 <td>BM906296</td> </td>	10 <td>BM906296</td>	BM906296
25	533 <td>72.0 <td>615 <td>9 <td>BM1346818</td> </td></td></td>	72.0 <td>615 <td>9 <td>BM1346818</td> </td></td>	615 <td>9 <td>BM1346818</td> </td>	9 <td>BM1346818</td>	BM1346818
26	511.4 <td>69.1 <td>542 <td>12 <td>BM045282</td> </td></td></td>	69.1 <td>542 <td>12 <td>BM045282</td> </td></td>	542 <td>12 <td>BM045282</td> </td>	12 <td>BM045282</td>	BM045282
27	488.4 <td>66.0 <td>518 <td>13 <td>BM626120</td> </td></td></td>	66.0 <td>518 <td>13 <td>BM626120</td> </td></td>	518 <td>13 <td>BM626120</td> </td>	13 <td>BM626120</td>	BM626120
28	477.4 <td>64.5 <td>567 <td>9 <td>BM528254</td> </td></td></td>	64.5 <td>567 <td>9 <td>BM528254</td> </td></td>	567 <td>9 <td>BM528254</td> </td>	9 <td>BM528254</td>	BM528254
29	471 <td>63.6 <td>572 <td>14 <td>BM439478</td> </td></td></td>	63.6 <td>572 <td>14 <td>BM439478</td> </td></td>	572 <td>14 <td>BM439478</td> </td>	14 <td>BM439478</td>	BM439478
30	459.2 <td>63.4 <td>520 <td>12 <td>BM1468250</td> </td></td></td>	63.4 <td>520 <td>12 <td>BM1468250</td> </td></td>	520 <td>12 <td>BM1468250</td> </td>	12 <td>BM1468250</td>	BM1468250
31	468.4 <td>63.3 <td>486 <td>12 <td>BM012745</td> </td></td></td>	63.3 <td>486 <td>12 <td>BM012745</td> </td></td>	486 <td>12 <td>BM012745</td> </td>	12 <td>BM012745</td>	BM012745
32	463 <td>62.6 <td>500 <td>10 <td>BM973297</td> </td></td></td>	62.6 <td>500 <td>10 <td>BM973297</td> </td></td>	500 <td>10 <td>BM973297</td> </td>	10 <td>BM973297</td>	BM973297
33	450.4 <td>60.9 <td>481 <td>9 <td>BM4573948</td> </td></td></td>	60.9 <td>481 <td>9 <td>BM4573948</td> </td></td>	481 <td>9 <td>BM4573948</td> </td>	9 <td>BM4573948</td>	BM4573948
34	446 <td>60.3 <td>446 <td>12 <td>BM726665</td> </td></td></td>	60.3 <td>446 <td>12 <td>BM726665</td> </td></td>	446 <td>12 <td>BM726665</td> </td>	12 <td>BM726665</td>	BM726665
35	441.8 <td>59.7 <td>465 <td>9 <td>BM582937</td> </td></td></td>	59.7 <td>465 <td>9 <td>BM582937</td> </td></td>	465 <td>9 <td>BM582937</td> </td>	9 <td>BM582937</td>	BM582937
36	431.6 <td>58.3 <td>700 <td>14 <td>BM429243</td> </td></td></td>	58.3 <td>700 <td>14 <td>BM429243</td> </td></td>	700 <td>14 <td>BM429243</td> </td>	14 <td>BM429243</td>	BM429243
37	430.4 <td>58.2 <td>698 <td>14 <td>BM452677</td> </td></td></td>	58.2 <td>698 <td>14 <td>BM452677</td> </td></td>	698 <td>14 <td>BM452677</td> </td>	14 <td>BM452677</td>	BM452677
38	426 <td>57.6 <td>695 <td>14 <td>BM422326</td> </td></td></td>	57.6 <td>695 <td>14 <td>BM422326</td> </td></td>	695 <td>14 <td>BM422326</td> </td>	14 <td>BM422326</td>	BM422326
39	420.6 <td>56.8 <td>1093 <td>12 <td>BM920038</td> </td></td></td>	56.8 <td>1093 <td>12 <td>BM920038</td> </td></td>	1093 <td>12 <td>BM920038</td> </td>	12 <td>BM920038</td>	BM920038
40	409.8 <td>55.4 <td>421 <td>9 <td>BM278825</td> </td></td></td>	55.4 <td>421 <td>9 <td>BM278825</td> </td></td>	421 <td>9 <td>BM278825</td> </td>	9 <td>BM278825</td>	BM278825
41	408 <td>55.1 <td>671 <td>14 <td>BM421971</td> </td></td></td>	55.1 <td>671 <td>14 <td>BM421971</td> </td></td>	671 <td>14 <td>BM421971</td> </td>	14 <td>BM421971</td>	BM421971
42	407 <td>55.0 <td>669 <td>14 <td>BM429601</td> </td></td></td>	55.0 <td>669 <td>14 <td>BM429601</td> </td></td>	669 <td>14 <td>BM429601</td> </td>	14 <td>BM429601</td>	BM429601
43	390 <td>52.7 <td>565 <td>12 <td>BM62328</td> </td></td></td>	52.7 <td>565 <td>12 <td>BM62328</td> </td></td>	565 <td>12 <td>BM62328</td> </td>	12 <td>BM62328</td>	BM62328
44	385.4 <td>52.1 <td>583 <td>14 <td>BM436146</td> </td></td></td>	52.1 <td>583 <td>14 <td>BM436146</td> </td></td>	583 <td>14 <td>BM436146</td> </td>	14 <td>BM436146</td>	BM436146
45	382 <td>51.6 <td>395 <td>9 <td>BM1262374</td> </td></td></td>	51.6 <td>395 <td>9 <td>BM1262374</td> </td></td>	395 <td>9 <td>BM1262374</td> </td>	9 <td>BM1262374</td>	BM1262374

ALIGNMENTS

RESULT 1
BM459678
LOCUS
DEFINITION
AGENCOURT_6417789 NIH_MGC_71 Homo sapiens cdna clone IMAGE:5534308
5', mRNA sequence.
ACCESSION
BM459678
VERSION
BM459678.1 GI:18508718
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 930)
NIH-MGC http://mgs.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1220 row: j column: 05
High quality sequence stop: 670.

FEATURES		Location/Qualifiers					
	source	1.	.930				
		/organism="Homo sapiens"					
		/mol_type="mRNA"					
		/db_xref="taxon:9606"					
		/clone="IMAGE:5534308"					
		/tissue_type="leiomyosarcoma"					
		/lab_host="DH10B (phage-resistant)"					
		/clone_id="NH_MGC_71"					
		/note="Organ: uterus; Vector: pCMV-Sport6, Site 1: NotI, Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt."					
		Average insert size 2.1 kb.					
BASE COUNT		182 a	310 c	263 g	172 t	3 others	
ORIGIN							
Query Match		98.6%	Score 729.4,	DB 12,	Length 930;		
Best Local Similarity		99.7%;	Pred. No. 8.5e-154;				
Matches 730;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;		
OY	9	GCTGACAGGGGGGTGAGCCGCCCGCAGACGGCCCGGGGATCTCTGAGACCATTGGGCGCGCCGCC	68				
Dd	1	GCTGACAGGGGGGTGAGCCGCCCGCAGACGGCCCGGGGATCTCTGAGACCATTGGGCGCGCCGCC	60				
OY	69	GACGCCATCTCTGATGTCGCTGGAGAAGTCTGACCCGCGAGAGACTCAAGAGTTCAAAGCTG	128				
Dd	61	GACGCCATCTCTGATGTCGCTGGAGAAGTCTGACCCGCGAGAGACTCAAGAGTTCAAAGCTG	120				
OY	129	AAGCTGCTGTGCGTGCCGCTGCGAGAGGCTTAAGGCGCATCCCGCGGCGCGCGTGTGCTG	188				
Dd	121	AAGCTGCTGTGCGTGCCGCTGCGAGAGGCTTAAGGCGCATCCCGCGGCGCGCGTGTGCTG	180				
OY	189	TCCATGAGACGCTTGGAGCTTACCCGACAAGCTGTGACGTTTAACTTGAAGCTTAAGGCT	248				
Dd	181	TCCATGAGACGCTTGGAGCTTACCCGACAAGCTGTGACGTTTAACTTGAAGCTTAAGGCT	240				
OY	249	GCCGAGCTCACCGCTAACGTGTGTCGCGCAGATGGGCTGCGAGAGATGGCGGGGAGAGCTG	308				
Dd	241	GCCGAGCTCACCGCTAACGTGTGTCGCGCAGATGGGCTGCGAGAGATGGCGGGGAGAGCTG	300				
OY	309	CAGGCGGCGACGACCAAGAGGCTCTTGAGACCGCGCCAGCTGGAGTCCAGGCCCCCTCTCTAG	368				
Dd	301	CAGGCGGCGACGACCAAGAGGCTCTTGAGACCGCGCCAGCTGGAGTCCAGGCCCCCTCTCTAG	360				
OY	369	TGCGGAGCGCAAGCCAGAGGCTCTGCACTTTATAGACACGACCCGSGCTGCGCTTATGCGGAGG	428				
Dd	361	TGCGGAGCGCAAGCCAGAGGCTCTGCACTTTATAGACACGACCCGSGCTGCGCTTATGCGGAGG	420				
OY	429	GTCACAAACGTTGAGTAGTGCTGTGATGTCGTACGCGGAAAGTCTTACCGAATGAGAG	488				
Dd	421	GTCACAAACGTTGAGTAGTGCTGTGATGTCGTACGCGGAAAGTCTTACCGAATGAGAG	480				
OY	489	TACGAGGAGTGGCGGCGGAGCCACCAACCCAAGAGTGGAGAGCTCTTCAAGTTTC	548				
Dd	481	TACGAGGAGTGGCGGCGGAGCCACCAACCCAAGAGTGGAGAGCTCTTCAAGTTTC	540				
OY	549	AACACGAGCTTGAAGCTGCAAGGACTTGTCTTCCAGAGCCCTTAAGGAGTCCAG	608				
Dd	541	AACACGAGCTTGAAGCTGCAAGGACTTGTCTTCCAGAGCCCTTAAGGAGTCCAG	600				
OY	609	TCTTACCTGTGGAGAGACTTGGAGGGGAGCTGAGGCTCTTCCAGCAACAATCCGGGTCA	668				
Dd	601	TCTTACCTGTGGAGAGACTTGGAGGGGAGCTGAGGCTCTTCCAGCAACAATCCGGGTCA	660				
OY	669	GCCCCTGGCAATCCAGCAAAATCATCTGATCTTTTATATACAAATATACGAAA	728				
Dd	661	GCCCCTGGCAATCCAGCAAAATCATCTGATCTTTTATATACAAATATACGAAA	720				
OY	729	AGCCAGCTTGAA	740				
Dd	721	AGCCAGCTTGAA	732				

LOCUS	BM926438	804 bp	mRNA	linear	EST 12-MAR-2002
DEFINITION	AGENCOURT_6644759 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766806				
ACCESSION	BM926438				
VERSION	5, mRNA sequence.				
KEYWORDS	BM926438.1 GI:19376817				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 804)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabs-remail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	CDNA Library Preparation: Life Technologies, Inc.				
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: LLM2824 row: a column: 15				
	High quality sequence stop: 746.				
FEATURES	Location/Qualifiers				
source	1..804				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5766806"				
	/lab_host="DH10B"				
	/clone_1b="NIH MGC 122"				
	/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."				
BASE COUNT	165 a 238 c 243 g 125 t 33 others				
ORIGIN					
Query Match	98.2%; Score 727; DB 12; Length 804;				
Best Local Similarity	99.2%; Pred. No. 2,96-153;				
Matches 730; Conservative	0; Mismatches 6; Indels 0; Gaps 0;				
0y	5 TCCGGCTGCAGCGGGGTGAGCGGCGGAGCGGCGGGATCTCTGAGCCATGTGGCGCGC 64				
Db	7 TCCCGGGATGACGGGGGTGAGCGGCGGCGGCGGCGGGGATCTGAGCCATGTGGCGCGC 66				
0y	65 GCGGACGCCATCTGTGATGCGCTGAGAACTTGACCGCGAGAGCTCAAGAACTTCAA 124				
Db	67 GCGGACGCCATCTGTGATGCGCTGAGAACTTGACCGCGAGAGCTCAAGAACTTCAA 126				
0y	125 GCTAAAGCTGTTCGTGCGTGCCTGCGCGGAGGCTTAAGGGCGATCCCCGGGGCGGCT 184				
Db	127 GCTAAAGCTGTTCGTGCGTGCCTGCGCGGAGGCTTAAGGGCGATCCCCGGGGCGGCT 186				
0y	185 GCTGTCATGAGCGCTTGAAGCTCAAGCAAGAGCTGTCAGCTTCTTACCTGAGAGCTTA 244				
Db	187 GCTGTCATGAGCGCTTGAAGCTCAAGCAAGAGCTGTCAGCTTCTTACCTGAGAGCTTA 246				
0y	245 CGGCGCGAGCTCAAGCTTAAGCTGCGCGAGCATGGGCTTGAGAGATGCGCGGCA 304				
Db	247 CGGCGCGAGCTCAAGCTTAAGCTGCGCGAGCATGGGCTTGAGAGATGCGCGGCA 306				
0y	305 GCTGACGGCGCGACCAAGAGCTTGAAGCGCGCCAGCTGGATCCAGGCGCCCTCC 364				


```

Db      307 GCTCAGGCGGCAACGACGACGAGGCTTGAGCCGCGCAGCTGGGATCCAGGCCCTCC 366
QY      365 TCAATCGGCAACCCAGACGAGGCTTGCACTTTATAGACCAACGAGGCTGCGTTATCGC 424
Db      367 TCAATCGGCAACCCAGACGAGGCTTGCACTTTATAGACCAACGAGGCTGCGTTATCGC 426
QY      425 GAGGGTCAACAACCTTGATGAGTGGCTGCTGATGCTCTGTAACGGAAGGCTTGAGCGGATGA 484
Db      427 GAGGGTCAACAACCTTGATGAGTGGCTGCTGATGCTCTGTAACGGAAGGCTTGAGCGGATGA 486
QY      485 GCAGTACCAAGGCACTGCGGCGCGAGCCCAACCAACCAAGATGCGGAAGCTCTTCAG 544
Db      487 GCAGTACCAAGGCACTGCGGCGCGAGCCCAACCAACCAAGATGCGGAAGCTCTTCAG 546
QY      545 TTTTACACCAAGCTTGGAATCTGGAATCTGTAAGGATCTTCTCCAGGCTCTTAAGGAGTC 604
Db      547 TTTTACACCAAGCTTGGAATCTGGAATCTGTAAGGATCTTCTCCAGGCTCTTAAGGAGTC 606
QY      605 CCAATCTTACCTGCTGAGGACCTGGAACGAGGCTGAGGCTCTTCCAGCAACATCCG 664
Db      607 CCAATCTTACCTGCTGAGGACCTGGAACGAGGCTGAGGCTCTTCCAGCAACATCCG 666
QY      665 GTGAGCCCTTGCAATCCCAACCAATCATCTGTAATCTGATCTTTTATACATATATAC 724
Db      667 GTGAGCCCTTGCAATCCCAACCAATCATCTGTAATCTGATCTTTTATACATATATAC 726
QY      725 GAAAGCCAGCTTGAA 740
Db      727 GAAAGCCAGCTTGAA 742

```

```

RESULT 3
LOCUS   BM910918
DEFINITION
  AGENCOURT 6615922 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:545497
  5', mRNA sequence.
ACCESSION
  BM910918
VERSION
  BM910918.1 GI:19361297
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Organism
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-remail.nih.gov
  Tissue Procurement: ATCC
  CDNA Library Preparation: Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LNL at:
  http://image.llnl.gov
  Plate: LNCM1949 row: 1 column: 18
  High quality sequence stop: 641.
  Location/Qualifiers
    1..767
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:545497"
      /rname="vastrocytoma grade IV, cell line"
      /lab host="DH10B (phage-resistant)"
      /clone_1lb="NIH_MGC_98"
      /note="Organ: brain; Vector: pOTB1; Site 1: XhoI; Site 2:
      EcoRI; cDNA made by oligo-dT priming. Directionally
      cloned into EcoRI/XhoI sites using the following 5'
      adaptor: GGCACGAG(G). Library constructed by Ling Hong
      in the laboratory of Gerald M. Rubin (University of
      California, Berkeley) using ZAP-cDNA synthesis kit

```

```

FEATURES
  source
    1..767
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:545497"
      /rname="vastrocytoma grade IV, cell line"
      /lab host="DH10B (phage-resistant)"
      /clone_1lb="NIH_MGC_98"
      /note="Organ: brain; Vector: pOTB1; Site 1: XhoI; Site 2:
      EcoRI; cDNA made by oligo-dT priming. Directionally
      cloned into EcoRI/XhoI sites using the following 5'
      adaptor: GGCACGAG(G). Library constructed by Ling Hong
      in the laboratory of Gerald M. Rubin (University of
      California, Berkeley) using ZAP-cDNA synthesis kit

```

```

BASE COUNT      163 a      241 c      237 g      124 t
ORIGIN

```

```

Query Match
Best Local Similarity 99.7%; Score 722; DB 12; Length 767;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY      7 CGGCTGACGCGGGGTGAGCGGCGGACGCGGCGGGGATCTTGAGCCATGAGGCGCGCGC 66
Db      17 CGGCTGACGCGGGGTGAGCGGCGGACGCGGCGGGGATCTTGAGCCATGAGGCGCGCGC 76
QY      67 GCGACGCCATCTCTGATAGCGCTGGAAGAACTGACCCGCGAGAGCTCAAGAGTTCAAGC 126
Db      77 GCGACGCCATCTCTGATAGCGCTGGAAGAACTGACCCGCGAGAGCTCAAGAGTTCAAGC 136
QY      127 TGAAGCTGCTGCGGATCCGCTGCGCGGAGGAGCTGAGGCGCATCCCGCGGGGCGCGCTGC 186
Db      137 TGAAGCTGCTGCTGCGGATCCGCTGCGCGGAGGAGCTGAGGCGCATCCCGCGGGGCGCGCTGC 196
QY      187 TGTTCATGAGACGCTTGGAACCTCAACGCAAGCTGATCAAGCTTCTTACCTGAGACCTACG 246
Db      197 TGTTCATGAGACGCTTGGAACCTCAACGCAAGCTGATCAAGCTTCTTACCTGAGACCTACG 256
QY      247 GCGCCGAGCTCACCGCTTAACGCTGCTGCGGACATGAGGCTTGCGAGAGATGCGCGGACG 306
Db      257 GCGCCGAGCTCACCGCTTAACGCTGCTGCGGACATGAGGCTTGCGAGAGATGCGCGGACG 316
QY      307 TGCAGGCGGCGCGACGACGACGAGGCTGAGGCGCGGAGGAGCTGACGAGGCGGCTCCTC 366
Db      317 TGCAGGCGGCGCGACGACGAGGCTGAGGCGCGGAGGAGCTGACGAGGCGGCTCCTC 376
QY      367 AGTCGAGGCGCAAGCCAGGCGCTGACCTTATAGACGACGACGCGGCTGCGCTTATCGCGA 426
Db      377 AGTCGAGGCGCAAGCCAGGCGCTGACCTTATAGACGACGACGCGGCTGCGCTTATCGCGA 436
QY      427 GGGTCACAAACGTTGATGCTGCTGATGCTGCTGTAACGGAAGGCTCTTACGAGATGAGC 486
Db      437 GGGTCACAAACGTTGATGCTGCTGATGCTGCTGTAACGGAAGGCTCTTACGAGATGAGC 496
QY      487 AGTACGAGGAGTGGCGGCGGAGCCGACCAACCAACCAAGATGCGGAGCTCTTACGTT 546
Db      497 AGTACGAGGAGTGGCGGCGGAGCCGACCAACCAACCAAGATGCGGAGCTCTTACGTT 556
QY      547 TCAACACCAAGCTGGAATCTGACCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGAGTCCC 606
Db      557 TCAACACCAAGCTGGAATCTGACCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGAGTCCC 616
QY      607 AGTCTTACTGCTGAGAGGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGT 666
Db      617 AGTCTTACTGCTGAGAGGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGT 676
QY      667 CAGGCGCTGCAATCCCAACCAATCATCTTGAATCTGATCTTTTATACATATATACG 725
Db      677 CAGGCGCTGCAATCCCAACCAATCATCTTGAATCTGATCTTTTATACATATATACG 736
QY      726 AAAAGCCAGCTTGAA 740
Db      737 AAAAGCCAGCTTGAA 751

```

```

RESULT 4
LOCUS   BM549665
DEFINITION
  AGENCOURT 6544142 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745741
  5', mRNA sequence.
ACCESSION
  BM549665
VERSION
  BM549665.1 GI:18785261
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Organism
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.
1 (bases 1 to 792)
AUTHORS
NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.

Mammalia; eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 792)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM12769 row: c column: 22
High quality sequence stop: 570.

FEATURES	Location/Qualifiers
source	1. .792

```

source
1. .792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5745741"
/cisuse_type="leukocyte"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_118"
/notes="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned. (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note
this is a NIH_MGC Library."
182 a 241 c 242 g 122 t 5 others
BASE COUNT
ORIGIN

```

Query Match	97.5%	Score 721.2	DB 12	Length 792
Best Local Similarity	98.8%	Pred. No. 5.7e-152		
Matches 726	Conservative 0	Mismatches 9	Indels 0	Gaps 0

QY	6	CCGGCTGACCGAGGGGTGAACCGGCGGACGGGCGGGAGATCTGGAGGCATATGGGGCGGCGG	65
Dp	19	CTCGATATCAGCGGGGTGAAGCGGCGGCGAGCGGCGGGAGATCTTGAAGCCATGGGGCGGCGG	78
QY	66	CGGACGCCCATCTGAGATGCGCTGAGAACCTGAACCGCGGAGAGCTCAAGAAATTCAAG	125
Dp	79	CGGAGCGGCATCTTGGAATGCGCTGGAGAACCTGAACCGCGGAGAGCTCAAGAAATTCAAG	138
QY	126	CTGAACTGCTGTCCGTGCGCGCTGGCGGAGGGCTTACGGGGCGCATCCGCGGGGCGCGCTG	185
Dp	139	CTGAACCTCTGTCCGTGCGCGCTGGCGGAGGGCTTACGGGGCGCATCCGCGGGGCGCGCTG	198
QY	186	CTGTTCATGAGCGCCTTGGACCTTCAACCGACAAAGCTGTGACCTTCTAACCTGGAGACCTAC	245
Dp	199	CTGTTCATGAGCGCCTTGGACCTTCAACCGACAAAGCTGTGACCTTCTAACCTGGAGACCTAC	258
QY	246	GGCGCGGAGCTTACCGGCTTACGTGCTGGCGGACATGGGCTTGACAGAGATGGCCGGGCGAG	305
Dp	259	GGCGCGGAGCTTACCGGCTTACGTGCTGGCGGACATGGGCTTGACAGAGATGGCCGGGCGAG	318
QY	306	CTGCGAGGGGSCACACGACCGAGGGGCTTGAAGCGGCGGACAGTGGAGATCCAGGCCCTTCT	365
Dp	319	CTGCGAGGGGSCACACGACCGAGGGGCTTGAAGCGGCGGACAGTGGAGATCCAGGCCCTTCT	378
QY	366	CAGTGGGACGCAAGCCAGGCTTGACCTTTATAGACGACACCGGGCTGGCTTATCGG	425
Dp	379	CAGTGGGACGCAAGCCAGGCTTGACCTTTATAGACGACACCGGGCTGGCTTATCGG	438
QY	426	AGGGTACAAACGTTGATGCTGCTGTGATGCTCTGTATCGGGAAGGTTCTGACGAGATGAG	485
Dp	439	AGGGTACAAACGTTGATGCTGCTGTGATGCTCTGTATCGGGAAGGTTCTGACGAGATGAG	498

QY	486	CAGTACCAAGGACAGTGCGGGCGAGGCCACCAACCAACGAAGTGCGAAGCTTTCAGT	545
Db	499	CAGTACCAAGGACAGTGCGGGCGAGGCCACCAACCAACGAATGCGAAGCTTTCAGT	558
QY	546	TTCAACAACAGCTTGGAATCGAAGGACTTGTCTCTCCAGAGCCCTTAAGGAGTCC	605
Db	559	TTCAACAACAGCTTGGAATCGAAGGACTTGTCTCTCCAGAGCCCTTAAGGAGTCC	618
QY	606	CAGTCTTAACCTGATGAGAGGACCTGAGAGGAGAGCTGAGGCTCTTCCACAGCAACCTCCGG	665
Db	619	CAGTCTTAACCTGATGAGAGGACCTGAGAGGAGAGCTGAGGCTCTTCCACAGCAACCTCCGG	678
QY	666	TCAGCCCCCTGGCAATCCCAACCAATTCATCTGAACTCTTTTATACACAATATACG	725
Db	679	TCAGCCCCCTGGCAATCCCAACCAATTCATCTGAACTCTTTTATACACAATATACG	738
QY	726	AAAGCCAGCTTGAA	740
Db	739	AAAGCCCGAATAA	753

RESULT	5
BG764161	
LOCUS	BG764161
DEFINITION	750 bp mRNA linear EST 15-MAY-2001 602737190F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862182 5', mRNA sequence.
ACCESSION	BG764161
VERSION	BG764161
KEYWORDS	EST. GI:14074814
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 750)
NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph. D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
plate: LLCM722 row: d column: 23
high quality sequence stop: 721.

FEATURES
Source

```

/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
/clone="IMAGE:462182"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

Query Match	96.4%	Score 713.4	DB 12	Length 750
Best Local Similarity	99.7%	Pred. No. 3.2e-150		
Matches 725; Conservative	0	Mismatches 1	Indels 1	Gaps 1

QY 15 GCGGGGTGAGCGGCGGAGCGGCGGAGATCTGAGAGCCATGAGGCGCGCGCGAGCC 74
 DB 2 GCGGGGTGAGCGGCGGAGCGGCGGAGATCTGAGAGCCATGAGGCGCGCGCGAGCC 61
 QY 75 ATCTGTGATGCGGTGAGGAACTGACCGCGGAGAGTCAAGAAATTCAAGCTGAACCTG 134
 DB 62 ATCTGTGATGCGGTGAGGAACTGACCGCGGAGAGTCAAGAAATTCAAGCTGAACCTG 121
 QY 135 CTGTGCGTGTGCGGTGCGGAGGCTGACGAGGCGTATCCCGCGGCGCGCTGTCTCATG 194
 DB 122 CTGTGCGTGTGCGGTGCGGAGGCTGACGAGGCGTATCCCGCGGCGCGCTGTCTCATG 181
 QY 195 GAGCGCTTGAAGCTTCAACGAGAGTGTGACCTTCACTGAGAGCTTACGAGCGCGAG 254
 DB 182 GAGCGCTTGAAGCTTCAACGAGAGTGTGACCTTCACTGAGAGCTTACGAGCGCGAG 241
 QY 255 CTGACCGCTTAACTGCTGCGGAGATGAGGCTGAGAGATGAGCGGCGAGCTGACGAG 314
 DB 242 CTGACCGCTTAACTGCTGCGGAGATGAGGCTGAGAGATGAGCGGCGAGCTGACGAG 301
 QY 315 GCGACGCGACCGAGGCTTGTGAGCGCGCGAGCTGAGATGAGCGCGCTCTCTCACTGCGCA 374
 DB 302 GCGACGCGACCGAGGCTTGTGAGCGCGCGAGCTGAGATGAGCGCGCTCTCTCACTGCGCA 361
 QY 375 GCGACGCGACCGAGGCTTGTGATGACGAGACCGCGCGCTTATGCGAGGCTTCA 434
 DB 362 GCGACGCGACCGAGGCTTGTGATGACGAGACCGCGCGCTTATGCGAGGCTTCA 421
 QY 435 AACGTTGAGTGTGCTGAGATGCTGTGATGAGGAGGCTGAGAGTGTGAGAGTGTGAG 494
 DB 422 AACGTTGAGTGTGCTGAGATGCTGTGATGAGGAGGCTGAGAGTGTGAGAGTGTGAG 481
 QY 495 GCAAGTGTGAGGCTGAGAGGCGCGACCAACCAAGAGATGAGAGGCTTCACTTCACTGACCA 554
 DB 482 GCAAGTGTGAGGCTGAGAGGCGCGACCAACCAAGAGATGAGAGGCTTCACTTCACTGACCA 541
 QY 555 GCGTGTGAGTGTGAGAGGCTGAGAGGCTTCTCTCCAGGAGGCTTCACTTCACTGAGTGTGAG 614
 DB 542 GCGTGTGAGTGTGAGAGGCTGAGAGGCTTCTCTCCAGGAGGCTTCACTTCACTGAGTGTGAG 601
 QY 615 CTGTGTGAGAGGCTGAGAGGCTGAGAGGCTTCTCTCCAGGAGGCTTCACTTCACTGAGTGTGAG 674
 DB 602 CTGTGTGAGAGGCTGAGAGGCTGAGAGGCTTCTCTCCAGGAGGCTTCACTTCACTGAGTGTGAG 661
 QY 675 GCGAATCCCAACCAATCATCTGATCTGATCTTTTATATACA-CAATATATGAGAAAGCA 733
 DB 662 GCGAATCCCAACCAATCATCTGATCTGATCTTTTATATAGAGAAATATGAGAAAGCA 721
 QY 734 GCTTGA 740
 DB 722 GCTTGA 728

RESULT 6
 CA309883/c
 LOCUS CA309883 725 bp mRNA linear EST 01-NOV-2002
 DEFINITION UI-H-FT1-bid-e-07-0-UI.s1 NCI CGAP FT1 Homo sapiens cDNA clone
 CA309883
 ACCESSION UI-H-FT1-bid-e-07-0-UI 3', mRNA sequence.
 VERSION CA309883.1 GI:24472937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: ggapbs-remail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I

FEATURES
 source
 location/Qualifiers
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FT1-bid-e-07-0-UI"
 /tissue_type="Alveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_1b="NCI CGAP FT1"
 /note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site_1: BclR 1; Site_2: Not I; NCI CGAP FT1 is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an BclR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCATGCGG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
 TAG LIB=UI-H-FT1
 TAG TISSUE=Human Lung Alveolar Macrophage
 TAG_SBO=GGCATGCGG"

BASE COUNT 119 a 221 c 226 g 159 t
 ORIGIN
 Query Match 95.24; Score 704.8; DB 14; Length 725;
 Best Local Similarity 99.74; Pred. No. 2,7e-148;
 Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 GCGGCGGAGATCTGAGAGGCTGAGGCGCGCGAGCCATCTGATGCGCTGAG 92
 DB 724 GCGGCGGAGATCTGAGAGGCTGAGGCGCGCGAGCCATCTGATGCGCTGAG 665
 QY 93 AACCTGACCGCGAGAGGCTCAAGAAATTCAAGTGAAGTGTGTGCGGCTGCGC 152
 DB 664 AACCTGACCGCGAGAGGCTCAAGAAATTCAAGTGAAGTGTGTGCGGCTGCGC 605
 QY 153 GAGGCGTACGAGGCGCATCCCGCGGCGCGCTGCTGATGAGAGGCTTGAACCTCAAC 212
 DB 604 GAGGCGTACGAGGCGCATCCCGCGGCGCGCTGCTGATGAGAGGCTTGAACCTCAAC 545
 QY 213 GACAGTGTGATGATCTTACTGAGAGCTTACGCGCGCGAGCTCACGCTTAACGTCTG 272
 DB 544 GACAGTGTGATGATCTTACTGAGAGCTTACGCGCGCGAGCTCACGCTTAACGTCTG 485
 QY 273 CGGCACTGAGGCTTGAAGAGATGCGCGGCACTGAGAGGCGCGCAACGAGGCTCT 332
 DB 484 CGGCACTGAGGCTTGAAGAGATGCGCGGCACTGAGAGGCGCGCAACGAGGCTCT 425
 QY 333 GAGGCGCGCGAGCTGAGATCCAGGCGCGCTCTCACTGAGAGGCGCGCAACGAGGCTCTCA 392
 DB 424 GAGGCGCGCGAGCTGAGATCCAGGCGCGCTCTCACTGAGAGGCGCGCAACGAGGCTCTCA 365
 QY 393 TTTATAGACCAAGACCGGCGTGTGCTTATGCGAGAGGCTCAAACTGTTAGTGTGCTCTG 452
 DB 364 TTTATAGACCAAGACCGGCGTGTGCTTATGCGAGAGGCTCAAACTGTTAGTGTGCTCTG 305
 QY 453 GATGCTTGTACGAGGAGGCTCTGAGAGGATGAGACGATACCAAGCAAGTGTGCGGCGGAGCC 512

Db 304 GATGCTGTACGGGAAGTCTTACGAGTACGATACGATACGAGGATGCGGCGGAGCCC 245

Qy 513 ACCAACCCAGACAGATGCGGAAGCTCTTCACTTCAACCAAGCTGGAACCTGAC 572

Db 244 ACCAACCCAGACAGATGCGGAAGCTTTCAGTTTCAACCAAGCTGGAACCTGAC 185

Qy 573 AAGGACTGTGCTCCCTCCAGGCGCCCTAAGGAGGCGGCTTACCTGAGGAGAGCTGAG 632

Db 184 AAGGACTGTGCTCCCTCCAGGCGCCCTAAGGAGGCGGCTTACCTGAGGAGAGCTGAG 125

Qy 633 CGGAGCTGAGGCTCTCCAGCAACACTCGGTGAGCGCCCTGAGCAATCCCAATCA 692

Db 124 CGGAGCTGAGGCTCTCCAGCAACACTCGGTGAGCGCCCTGAGCAATCCCAATCA 65

Qy 693 TCCTGAATCTGATCTTTTATACCAATATACGAAAGCCAGCTTGA 740

Db 64 TCCTGAATCTGATCTTTTATACCAATATACGAAAGCCAGCTTGA 17

RESULT 7
Bg767422 856 bp mRNA linear EST 15-MAY-2001
LOCUS 602741219F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4871020 5',
DEFINITION mRNA sequence.
ACCESSION Bg767422
VERSION Bg767422.1 GI:14078075
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 856)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1745 row: e column: 05
High quality sequence stop: 696.
Location/Qualifiers
1. 856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4871020"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 194 a 234 c 279 g 149 t

ORIGIN

Query Match 93.4%; Score 690.8; DB 12; Length 856;
Best Local Similarity 99.6%; Pred. No. 4e-145;
Matches 703; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 15 GCGGGGTGAGCGGCGGCGGCGGCGGATCTGAGCGCATGGGCGGCGGCGGCGGCGG 74

Db 2 GCGGGGTGAGCGGCGGCGGCGGCGGCGGATCTGAGCGCATGGGCGGCGGCGGCGGCGG 61

Qy 75 ATCTGAGATGCGGTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGTGAAGCTG 134

Db 62 ATCTGAGATGCGGTGAGAACTGACCGCGGAGAGCTCAAGAGTTCAAGTGAAGCTG 121

Qy 135 CTGTGCGTGGCTGCGGCGGAGGCTTACCGGCGCATCCCGCGGCGGCGGCTGTCTCATG 194

Db 122 CTGTGCGTGGCTGCGGCGGAGGCTTACCGGCGCATCCCGGCGGCGGCGGCTGTCTCATG 181

Qy 195 GAGCGCTGAGCTCAACCGCAAGCTGATGAGCTTACCTGAGAGCTTACCGGCGGAG 254

Db 182 GAGCGCTGAGCTCAACCGCAAGCTGATGAGCTTACCTGAGAGCTTACCGGCGGAG 241

Qy 255 CTCACCGCTAACGCTGCGCGCATGAGGCTGCGAGAGATGCGCGGCGAGCTGACAGCG 314

Db 242 CTCACCGCTAACGCTGCGCGCATGAGGCTGCGAGAGATGCGCGGCGAGCTGACAGCG 301

Qy 315 GCCACGCAACGAGGCTCTGAGAGCGCGGCGGAGTCCAGGCCCTCTCACTGAGGCA 374

Db 302 GCCACGCAACGAGGCTCTGAGAGCGCGGCGGAGTCCAGGCCCTCTCACTGAGGCA 361

Qy 375 GCCACGCAACGAGGCTCTGAGAGCGCGGCGGAGTCCAGGCCCTCTCACTGAGGCA 434

Db 362 GCCACGCAACGAGGCTCTGAGAGCGCGGCGGAGTCCAGGCCCTCTCACTGAGGCA 421

Qy 435 AACGTTAGTGTGCTGTGATGCTGTATCGGAGAGTCTGACGATGACAGTACAG 494

Db 422 AACGTTAGTGTGCTGTGATGCTGTATCGGAGAGTCTGACGATGACAGTACAG 481

Qy 495 GCGATGCGGCGGAGCGGCGGAGCGGAGCGGAGAGTGGAGAGCTTCAAGTTTCAACCA 554

Db 482 GCGATGCGGCGGAGCGGCGGAGCGGAGCGGAGAGTGGAGAGCTTCAAGTTTCAACCA 541

Qy 555 GCCTGGAATCTGAGCTGCAAGAGCTTGTCTCAGAGCGCTTAAGGAGTCCAGTCTAC 614

Db 542 GCCTGGAATCTGAGCTGCAAGAGCTTGTCTCAGAGCGCTTAAGGAGTCCAGTCTAC 601

Qy 615 CTGTGAGAGAGCTGAGAGCGGAGCTGCTCTTCCAGCAACACTCGGTGAGCGGCT 674

Db 602 CTGTGAGAGAGCTGAGAGCGGAGCTGAGGCTCTTCCAGCAACACTCGGTGAGCGGCT 661

Qy 675 GGC-AATCCCAACCAATCATCTGATCTGATCTTTTATACCA 719

Db 662 GGCATATCCCAACCAATCATCTGATCTGATCTTTTATACCA 707

RESULT 8
Bg337806 748 bp mRNA linear EST 27-FEB-2001
LOCUS 602435709F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4553378 5',
DEFINITION mRNA sequence.
ACCESSION Bg337806
VERSION Bg337806.1 GI:13144342
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 748)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1251 row: b column: 03

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 676)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548

COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=yes

FEATURES
 source
 1..676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-abr-k-24-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /note="Organ: Lung; Vector: pRTT3-Pac (Pharmacia) with a
 modified polylinker; Site_1: Ecor I; Site_2: Not I;
 UI-CF-EC1 is a normalized CDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Ronaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA,
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into pRTT3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AAGTGGCTTAC.
 TAG_LIB=UI-CF-EC1
 TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG_SEQ=AACTGCTTAC"

BASE COUNT 112 a 199 c 211 g 154 t

ORIGIN

Query Match 89.2%; Score 660; DB 12; Length 676;
 Best Local Similarity 100.0%; Pred. No. 3,2e-138;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GATGCGCTGGAGAACTGACCGCGAGAGCTCAAGAAATTCAAGCTGAGCTGTGCT 140
 Db 676 GATGCGCTGGAGAACTGACCGCGAGAGCTCAAGAAATTCAAGCTGAGCTGTGCT 617
 QY 141 GTGCGCGTCCGCGGCGCTAGCGCGCATCCCGCGGCGCGCGCTGCTCATGAGACGCC 200
 Db 616 GTGCGCGTCCGCGGCGCTAGCGCGCATCCCGCGGCGCGCGCTGCTCATGAGACGCC 557
 QY 201 TTGAGACCTCAACGCAAGCTGATGCTTCTACCTGAGACCTCAACGCGCGGAGCTCAC 260
 Db 556 TTGAGACCTCAACGCAAGCTGATGCTTCTACCTGAGACCTCAACGCGCGGAGCTCAC 497

QY 261 GCTTACGCTGCTGGCGCCACATGCGCTTGCAGAGATGCGCGGCGAGCTGCGGCCACG 320
 Db 496 GCTTACGCTGCTGGCGCCACATGCGCTTGCAGAGATGCGCGGCGAGCTGCGGCCACG 437
 QY 321 CACGAGGCTCTGAGCGCGCGCCAGCTGGATCCAGGCCCTCTCTGATGCGGCGCCAG 380
 Db 436 CACGAGGCTCTGAGCGCGCGCCAGCTGGATCCAGGCCCTCTCTGATGCGGCGCCAG 377
 QY 381 CAGGCTGCACTTTATAGACAGACACCGGCTGCGCTTATCGGAGGTCACAAAGTT 440
 Db 376 CAGGCTGCACTTTATAGACAGACACCGGCTGCGCTTATCGGAGGTCACAAAGTT 317
 QY 441 GAGTGGCTGCTGATGCTCTGTATCGGAGAGTCTTGAACGATGACATACAGGCACTG 500
 Db 316 GAGTGGCTGCTGATGCTCTGTATCGGAGAGTCTTGAACGATGACATACAGGCACTG 257
 QY 501 CGGCGCGAGCCACCAACCCAGCAAGATCGGAGAGTCTTCAAGTTTCAACAGCCCTGG 560
 Db 256 CGGCGCGAGCCACCAACCCAGCAAGATCGGAGAGTCTTCAAGTTTCAACAGCCCTGG 197
 QY 561 AACTGACCTGCAAGAGCTTGTCTCTCCAGGCTTAAAGAGTCCAGTCCATCTGATG 620
 Db 196 AACTGACCTGCAAGAGCTTGTCTCTCCAGGCTTAAAGAGTCCAGTCCATCTGATG 137
 QY 621 GAGGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTAGCCCTGGCAAT 680
 Db 136 GAGGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTAGCCCTGGCAAT 77
 QY 681 CCCACCAATCATCTGATCTGATCTTTTATACAAATATGAAAGCAGCTTGA 740
 Db 76 CCCACCAATCATCTGATCTGATCTTTTATACAAATATGAAAGCAGCTTGA 17

RESULT 12
 BE908204
 LOCUS BE908204 675 bp mRNA linear EST 20-OCT-2000
 DEFINITION 601500461P1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902319 5',
 mRNA sequence.
 ACCESSION BE908204
 VERSION BE908204.1 GI:10402543
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 675)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strauberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: IMAGE704 row: j column: 16
 High quality sequence stop: 660.

FEATURES
 source
 1..675
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3902319"
 /tissue_type="epithelioid carcinoma"
 /note="Organ: pancreas; Vector: PCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life

BASE COUNT 151 a 214 c 198 g 112 t
 ORIGIN

Query Match 88.4%; Score 654.2; DB 10; Length 675;
 Best Local Similarity 99.5%; Pred. No. 6.5e-137;
 Matches 656; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

82 ATGCGCTGGAGAACTGACCGCGGAGAGCTCAAGAAATTCAAGACTGCTGTGG 141
 1 ATGCGCTGGAGAACTGACCGCGGAGAGCTCAAGAAATTCAAGACTGCTGTGG 60
 142 TGGCGCTGGCGGAGGCTACGGGCGCATCCCGGGGGGGGCTGCTGTCAATGAGCC 201
 61 TGGCGCTGGCGGAGGCTACGGGCGCATCCCGGGGGGGGCTGCTGTCAATGAGCC 120
 202 TGGAGCTCAACGCAAGCTGTGAGCTTCTACCTGGAGACCTTACGGCGCGGAGCTACCG 261
 121 TGGAGCTCAACGCAAGCTGTGAGCTTCTACCTGGAGACCTTACGGCGCGGAGCTACCG 180
 262 CTAACTGTCTGGCGGAGATGGGCTTGGAGAGATGGCGGGGAGCTTACGGCGCGGAG 321
 181 CTAACTGTCTGGCGGAGATGGGCTTGGAGAGATGGCGGGGAGCTTACGGCGCGGAG 240
 322 ACCAGGCTCTGGAGCGCGGCGGAGCTGAGATCCAGGCGCTCTCTCACTGGCGGAGCG 381
 241 ACCAGGCTCTGGAGCGCGGCGGAGCTGAGATCCAGGCGCTCTCTCACTGGCGGAGCG 300
 382 CAGGCTGTGACTTTATAGACCAAGCAAGGCTGCTTATCGGAGGCTCAAGAGCTTG 441
 301 CAGGCTGTGACTTTATAGACCAAGCAAGGCTGCTTATCGGAGGCTCAAGAGCTTG 360
 442 AGTGTCTGTGAGTGTCTGTATCGGAGAGGCTCTTACGATGAGACGATACGAGCACTG 501
 361 AGTGTCTGTGAGTGTCTGTATCGGAGAGGCTCTTACGATGAGACGATACGAGCACTG 420
 502 GGGCGGAGCGGCGGAGCAAGCAAGGAGATGGGAGGCTTCACTTCAACAGCGCTTGA 561
 421 GGGCGGAGCGGCGGAGCAAGCAAGGAGATGGGAGGCTTCACTTCAACAGCGCTTGA 480
 562 ACTGAGCTGCAAGAGCTTGTCTCTCCAGGCTTAAAGGAGTCCAGCTCTTAACTGATG 621
 481 ACTGAGCTGCAAGAGCTTGTCTCTCCAGGCTTAAAGGAGTCCAGCTCTTAACTGATG 540
 622 AGGAGCTGCAAGGAGCTGAGGCTCTTCCAGCAACACTCCGCTGAGCGCTTGAATC 681
 541 AGGAGCTGCAAGGAGCTGAGGCTCTTCCAGCAACACTCCGCTGAGCGCTTGAATC 600
 682 CCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGAAAGCGAGCTTGA 740
 601 CCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGAAAGCGAGTTGAA 659

RESULT 13
 BG255521 744 bp mRNA linear EST 13-FEB-2001
 LOCUS 602367671F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476088 5',
 DEFINITION mRNA sequence.
 ACCESSION BG255521
 VERSION BG255521.1 GI:12765259
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 744)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph. D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LIML10302 row: m column: 17
 High quality sequence stop: 661.
 Location/Qualifiers

FEATURES

source
 1..744
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4476088"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 91"
 /note="Organ: prostate; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."
 BASE COUNT 159 a 227 c 238 g 120 t

Query Match 87.5%; Score 647.4; DB 10; Length 744;
 Best Local Similarity 97.0%; Pred. No. 2.2e-135;
 Matches 713; Conservative 0; Mismatches 16; Indels 6; Gaps 5;

8 GGCTGAGGAGGAGTGAAGCGGCGGAGGAGTCTGAGAGCCATGAGGAGCGGCGG 67
 1 GGCTGAGGAGGAGTGAAGCGGCGGAGGAGTCTGAGAGCCATGAGGAGCGGCGG 60
 68 CAGCGCATCTGATGCTGCTGAGAGAACTGACCGCGGAGAGCTTCAAGATTCAAGCT 127
 61 CAGCGCATCTGATGCTGCTGAGAGAACTGACCGCGGAGAGCTTCAAGATTCAAGCT 120
 128 GAACTGCTGTGAGTGGCGGCTGCGGAGAGGCTTACGAGGCGGAGCGGCGGCTGCT 187
 121 GAACTGCTGTGAGTGGCGGCTGCGGAGAGGCTTACGAGGCGGAGCGGCGGCTGCT 180
 188 GTCAATGAGCGGCTTGAACCTTCAACGCAAGAGTGTGAGCTTCACTGAGAGCTTACG 247
 181 GTCAATGAGCGGCTTGAACCTTCAACGCAAGAGTGTGAGCTTCACTGAGAGCTTACG 240
 248 CGCGAGCTCAACGCTTAACTGCTGCGGAGCAATGAGGCTTGAAGAGATGCGGAGCT 307
 241 CGCGAGCTCAACGCTTAACTGCTGCGGAGCAATGAGGCTTGAAGAGATGCGGAGCT 300
 308 GCAAGGCGGAGGAGCAAGCAAGGCTTGAAGGCGGCGGAGATGCGGAGCTTCA 367
 301 GCAAGGCGGAGGAGCAAGCAAGGCTTGAAGGCGGCGGAGATGCGGAGCTTCA 360
 368 GTGCGAGGCAAGCGGAGGCTTGAATATAGACCAAGCAAGGAGTGTGAGGCTTATGCGAG 427
 361 GTGCGAGGCAAGCGGAGGCTTGAATATAGACCAAGCAAGGAGTGTGAGGCTTATGCGAG 420
 428 GGTCAAAAAGTTGAGTGTGCTGCTGTAACGGAAGTCTTGAAGAGATGAGCA 487
 421 GGTCAAAAAGTTGAGTGTGCTGCTGTAACGGAAGTCTTGAAGAGATGAGCA 480
 488 GTACCAAGGAGTGTGAGGCGGAGGCGGAGCCCAACCAAGCAAGATGCGGAGCTTCAAGTT 547
 481 GTACCAAGGAGTGTGAGGCGGAGGCGGAGCCCAACCAAGCAAGATGCGGAGCTTCAAGTT 539
 548 CACACAGGCTGGAAGTGAAGCTTGAAGAGATGCTTCCAGGCGCTTGAAGAGATGAGCA 607
 540 CACACAGGCTGGAAGTGAAGCTTGAAGAGATGCTTCCAGGCGCTTGAAGAGATGAGCA 597
 608 GTCTTACCTGTGAGAGAGCTTGAAGAGATGAGGCTTCTTCCAGCAAACTCCGATC 667
 598 GTCTTACCTGTGAGAGAGCTTGAAGAGATGAGGCTTCTTCCAGCAAACTCCGATC 656
 668 ACCCGCTGCAATCCCAACCAATCATCTGAT--CTGATCTTTTATACAAATATACG 725


```

Db      657 AGCCCGGGAAATCCACCAAAATATCTGATCTGATCTTTTATACCAATATACG 716
Qy      726 AAAAGCCAGCTTGAA 740
Db      717 AAAAGCCAGTTGAAA 731

RESULT 14
LOCUS   BE909218
DEFINITION 601501478F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903498 5',
mRNA sequence.
VERSION BE909218.1 GI:10404770
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 687)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Straubeberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM9707 row: K column: 19
High quality sequence stop: 673.
Location/Qualifiers
1..687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3903498"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/notes="Organ: pancreas; Vector: PCMV-SPORE; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 140 a 219 c 214 g 114 t
ORIGIN
Query Match 86.4%; Score 639; DB 10; Length 687;
Best Local Similarity 98.7%; Pred. No. 1.7e-113;
Matches 676; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

Qy      48 GAGGCCATGGGGGCGCGC-GCGCGACGCCATCTGTGATGCGCTGGAGAACCTGACCGCGA 106
Db      1 GAGGCAATGGGGGCGCGCGCGCGACGCCATCTGTGATGCGCTGGAGAACCTGACCGCGA 60

Qy      107 GAGGCTCAAAAGTTCAAGCTGAAGCTGCTGTGCGTGCCTGCGCGAGGCTTACGGGCG 166
Db      61 GAGGCTCAAAAGTTCAAGCTGAAGCTGCTGTGCGTGCCTGCGCGAGGCTTACGGGCG 120

Qy      167 CATCCGCGGGGGGCGCGCTGCTCATGAGAGCGCTTGAGACTCAACGAAAGCTGATCAG 226
Db      121 CATCCGCGGGGGGCGCGCTGCTCATGAGAGCGCTTGAGACTCAACGAAAGCTGATCAG 180

Qy      227 CTTTCTACTTGAGACTTACGCGCGCGAGCTTACCGCTTACGCTGCTGCGGAGCATGGGCT 286
Db      181 CTTTCTACTTGAGACTTACGCGCGCGAGCTTACCGCTTACGCTGCTGCGGAGCATGGGCT 240

Qy      287 GCAGGAGATGGCGGGGAGCTGACAGGGGCGCACGACCAAGGCTCTGGAGCCGGCGCAGC 346

```

```

Db      241 GCAGGAGATGGCGGGGAGCTGACAGCGCGCCACGACCAAGGGCTCTGGAGCCGCCAGC 300
Qy      347 TGGGATCCAGGCCCTCTCTCAGTCGCGAGGCCAAGCCAGGCTGCACTTTATAGACGACA 406
Db      301 TGGGATCCAGGCCCTCTCTCAGTCGCGAGGCCAAGCCAGGCTGCACTTTATAGACGACA 360
Qy      407 CCGGGCTGCGCTTATGCGAGGGTCACAAAAGTTAGTGTGCTGTGATGCTCTGTACGG 466
Db      361 CCGGGCTGCGCTTATGCGAGGGTCACAAAAGTTAGTGTGCTGTGATGCTCTGTACGG 420
Qy      467 GAAGTCTCGAGGATGAGACGATACGAGGAGTGGCGGGCGGAGGCCAACCAACCAACCA 526
Db      421 GAAGTCTCGAGGATGAGACGATACGAGGAGTGGCGGGCGGAGGCCAACCAACCAACCA 480
Qy      527 GATCGGAGACCTCTTCAAGTTTCAACACGAGCTGGAACCTGACCTGCAAGACTTGTCTCT 586
Db      481 GATCGGAGACCTCTTCAAGTTTCAACACGAGCTGGAACCTGACCTGCAAGACTTGTCTCT 540
Qy      587 CCAAGGCTTAAAGGAGTCCAGTCTTACTGTGTGAGAGACCTGAGCGGAGCTGAGGCTC 646
Db      541 CCAAGGCTTAAAGGAGTCCAGTCTTACTGTGTGAGAGACCTGAGCGGAGCTGAGGCTC 600
Qy      647 CTTCCAGCAACAACCTCGGTGACAGCCCTGGCAATCCACCAATCATCTGAATCTGATC 706
Db      601 -TTCCAGCAACAACCTCGGTGACAGCCCTGGG-AATCCACCAATCATCTGAATCTGATC 657
Qy      707 TTTTATACCAATATACGAAAAGC 731
Db      658 TTTTATACCAATATACGAAAAGC 682

```

```

RESULT 15
LOCUS   BM998354
DEFINITION 638 bp mRNA linear EST 17-JUN-2002
UI-H-DT1-awc-p-21-0-UI.61 NCI_CGAP_DTI Homo sapiens cDNA clone
IMAGE:5887748 3', mRNA sequence.
VERSION BM998354
KEYWORDS BM998354.1 GI:19723255
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 638)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished
Contact: Robert Straubeberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5887748"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="NCI_CGAP_DTI"
/lab_host="DH10B (Life Technologies)"

```

```

FEATURES
SOURCE
Note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DT1 is a normalized cDNA library containing the
following tissue(s): Metastatic Chondrosarcoma in Lung.
The library was constructed according to Bonaldo, Lemmon

```

and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTZ19-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tag for this library is AACTGTCCG.

BASE COUNT 108 a 185 c 201 g 144 t
ORIGIN
TAG_SEQ=AACTGTCCG"
TAG_TISSUE=lung metastatic chondrosarcoma

Query Match 83.4%; Score 617.2; DB 12; Length 638;
Best Local Similarity 99.5%; Pred. No. 1.3e-128;
Matches 619; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 119 GTTCAAGCTGAAGCTGCTGTCGCTGCGCGAGGGCTACGGGCGATCCCGCGGG 178
DB 638 GTCAAGCTGAAGCTGCTGTCGCTGCGCGAGGGCTACGGGCGATCCCGCGGG 579

QY 179 CGGCGTGTGTCATGAGAGCTTGAGCTTACCGAAGAGTGTGAGCTTCTTACTGGA 238
DB 578 CGGCGTGTGTCATGAGAGCTTGAGCTTACCGAAGAGTGTGAGCTTCTTACTGGA 519

QY 239 GACCTACGGCGCGAGCTTACCGCTTACGCTGCGCGAATGAGGCTGCGAGAGATGCG 298
DB 518 GACCTACGGCGCGAGCTTACCGCTTACGCTGCGCGAATGAGGCTGCGAGAGATGCG 459

QY 299 CGGCGAGCTGAGAGCGCGCGAGCTTGAAGCGCGCGAGCTGAGATCCAGGC 358
DB 458 CTGGCAGCTGAGAGCGCGCGAGCTTGAAGCGCGCGAGCTGAGATCCAGGC 399

QY 359 CCTCTCTAGTGGGCGAGCGAGCTTGAAGCTTGAAGCGAGAGTCTGCGCT 418
DB 398 CCTCTCTAGTGGGCGAGCGAGCTTGAAGCTTGAAGCGAGAGTCTGCGCT 339

QY 419 TATCGGAGGGTCAAAAGTGTGAGTGTGCTGTGATGCTGTGTACGGGAGTCTGTAC 478
DB 338 TATCGGAGGGTCAAAAGTGTGAGTGTGCTGTGATGCTGTGTACGGGAGTCTGTAC 279

QY 479 GATGAGCAGTACAGAGCAGTGTGCGCGCGAGCGCGAGCGAGAGTGTGAGAGCT 538
DB 278 GATGAGCAGTACAGAGCAGTGTGCGCGCGAGCGCGAGCGAGAGTGTGAGAGCT 219

QY 539 CTTCAAGTTTCAACAGAGCTGGAAGCTGGAAGCTTGTGCTCTCTCAAGGCTTAAAG 598
DB 218 CTTCAAGTTTCAACAGAGCTGGAAGCTGGAAGCTTGTGCTCTCTCAAGGCTTAAAG 159

QY 599 GAGTCCCAAGTCTTACCTGTGTGAGAGCTTGAAGCGAGCTGAGAGCTCTTCCAGCAAC 658
DB 158 GAGTCCCAAGTCTTACCTGTGTGAGAGCTTGAAGCGAGCTGAGAGCTCTTCCAGCAAC 99

QY 659 ACTCCGCTCAGCGCTTGGCAATCCCAAGCAATATCTGAATGTGATCTTTTATACCA 718
DB 98 ACTCCGCTCAGCGCTTGGCAATCCCAAGCAATATCTGAATGTGATCTTTTATACCA 39

QY 719 ATATAGAAAGCAGCTTGA 740
DB 38 ATATAGAAAGCAGCTTGA 17

Search completed: January 29, 2004, 13:05:55
Job time : 1480.73 secs

Db	241	CTTATGGGGCGAGGCTCACCGCTTAAGTCTGTGGCACTAGGGCTCTGCAGAGATGCGG	300
Qy	301	GGCAGCTGCAGGCGGCGCAACGCAACAAGGCTCTGAGCCGCGCCAGCTGGGATCCAGAGCC	360
Db	301	GGCAGCTCAGGCGGCGCGCACAGAGGCTCTGAGCCGCGCCAGCTGGGATCCAGAGCC	360
Qy	361	CTCCCTCAGTCGCGAGCCCAAGCCAGGCTGTGACTTTATACAGACACCGGCGGTGCGCTTA	420
Db	361	CTCTCTCAGTCGCGAGCCCAAGCCAGGCTGTGACTTTATACAGACACCGGCGGTGCGCTTA	420
Qy	421	TCGCGAGGGTCACAAAAGTTGAGTGGCTGTGATGCTCTGTACGGGAAGTCTTGACGG	480
Db	421	TCGCGAGGGTCACAAAAGTTGAGTGGCTGTGATGCTCTGTACGGGAAGTCTTGACGG	480
Qy	481	ATGAGCAATACCAAGGCATGCGGGCGAGGCCCAACCAACCAAGCAAGATGCGGAAGCTCT	540
Db	481	ATGAGCAATACCAAGGCATGCGGGCGAGGCCCAACCAACCAAGCAAGATGCGGAAGCTCT	540
Qy	541	TCAGTTTCACACCAAGCTTGAACTGGAAGACTTGTCTCTCCAGAGCCTTAAGGG	600
Db	541	TCAGTTTCACACCAAGCTTGAACTGGAAGACTTGTCTCTCCAGAGCCTTAAGGG	600
Qy	601	AGTCCCAATCCATCCTGTGTGAGAGACCTGGAGCGAGAGCTGAGGCTCTTCCAGCAACAC	660
Db	601	AGTCCCAATCCATCCTGTGTGAGAGACCTGGAGCGAGAGCTGAGGCTCTTCCAGCAACAC	660
Qy	661	TCGGGTACAGCCCTGGCAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT	720
Db	661	TCGGGTACAGCCCTGGCAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT	720
Qy	721	ATAGCAAAAGCAGGTTGAA	740
Db	721	ATAGCAAAAGCAGGTTGAA	740

```

RESULT 2
US-09-340-620A-50
; Sequence 50, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-50

```

	Query Match	Similarity	Score	DB	Length	Mismatches	Conservative	Indels	Gaps
	79.1%	100.0%	585	4	585	0	0	0	0
	Beet	Local							
	Matches	585							
QY	54	ATGGGGCGCGCGCGAGCCATCCTGGATGCGTGAAGACCTACCGCGAGGAGCTC	113						
DB	1	ATGGGGCGCGCGCGAGCCATCCTGGATGCGCTGAAGACTTACCGCGCAGGAGCTC	60						
QY	114	AAGAATTCAGACTGAGCTGCTGCGGACCGCGCGGAGAGGGCTACGGGCGCATCCG	173						
	61	AAGAATTCAGACTGAGCTGCTGCGGACCGCGCGGAGAGGGCTACGGGCGCATCCG	120						

QY	174	CGGGGCGCGCTGCTGTCCATGAGCGCCTTGGACCTCACCGACAAGCTGGTCAAGCTTCTAC	233
Db	121	CGGGGCGCGCTGCTGTCCATGAGCGCCTTGGACCTCACCGACAAGCTGGTCAAGCTTCTAC	180
QY	234	CTGGAGACCTTACGGGGCGGAGCTCAACCGCTTAAACGTCGCGCGAGATGGGCGCTGGACGAG	293
Db	181	CTGGAGACCTTACGGGGCGGAGCTCAACCGCTTAAACGTCGCGCGAGATGGGCGCTGGACGAG	240
QY	294	ATGACCGGGGACGTTGACAGGCGGCGCACCGACCGAGGCTCTGGAGCGCGCAACTGGGATC	353
Db	241	ATGACCGGGGACGTTGACAGGCGGCGCACCGACCGAGGCTCTGGAGCGCGCAACTGGGATC	300
QY	354	CAGGCGCCCTCCTCAAGTGGGCGACCAAGCGCAGGCGTGGCACTTTATGACAGGACCGGGCT	413
Db	301	CAGGCGCCCTCCTCAAGTGGGCGACCAAGCGCAGGCGTGGCACTTTATGACAGGACCGGGCT	360
QY	414	GCGGCTTATCGGAGGGGTACAAACGTTGAGTGGCTGTGGATGCTCTGTAACGGGAAGTTC	473
Db	361	GCGGCTTATCGGAGGGGTACAAACGTTGAGTGGCTGTGGATGCTCTGTAACGGGAAGTTC	420
QY	474	CTGACGGATGAGCAAGTACCAAGCAAGTGGGGGCGGAGCCCAACCAACCAAGCAAGATGGCG	533
Db	421	CTGACGGATGAGCAAGTACCAAGCAAGTGGGGGCGGAGCCCAACCAACCAAGCAAGATGGCG	480
QY	534	AAGCTCTTCAATTTACACACAGACCTGGAACTGACCTTGCAAGGACTTGTCTCTCCAGGCC	593
Db	481	AAGCTCTTCAATTTACACACAGACCTGGAACTGACCTTGCAAGGACTTGTCTCTCCAGGCC	540
QY	594	CTAAGGGAAGTCCCAAGTCTTACTGTGGAGGAACCTGGAGCGGAGC 638	
Db	541	CTAAGGGAAGTCCCAAGTCTTACTGTGGAGGAACCTGGAGCGGAGC 585	

```

RESULT 3
US-09-340-620A-60
: Sequence 60, Application US/09340620A
: Patent No. 6482933
: GENERAL INFORMATION
: APPLICANT: Bertlin, John
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
: FILE REFERENCE: 07334-124001
: CURRENT APPLICATION NUMBER: US/09/340,620A
: PRIOR FILING DATE: 1999-06-28
: PRIOR APPLICATION NUMBER: US 09/245,281
: PRIOR FILING DATE: 1999-02-05
: PRIOR APPLICATION NUMBER: US 09/207,359
: PRIOR FILING DATE: 1998-12-08
: PRIOR APPLICATION NUMBER: US 09/099,041
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: US 09/019,942
: PRIOR FILING DATE: 1998-02-06
: NUMBER OF SEQ ID NOS: 71
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 60
: LENGTH: 777
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (89)...(667)
: US-09-340-620A-60

```

	Query Match	45.4%	Score 3361	DB 4	Length 777
	Best Local Similarity	72.9%	Pred. No. 7.4e-67		
	Matches 461	Conservative	0	Mismatches 165	Indels 6
					Gaps 2
Oy	18	GGGTGACCGCGCGCAGCGCGCGCGGAGATCTTGAGACCATGCGCGCGCGCGGACGCCATTC	77		
Dh	53	GAGTAAAAAGGTGACCGCGGCGTGGCCACCCAGAGCAGATGGCGCGCGGACGAGATGCCATTC	112		
Oy	78	CTGATGCGCTGGAGAACTTGACCGCGGAGAGCTCAAGAAATTCAAGCTGAAGTCTGTC	137		

Db 113 CTGACGCTCTTGAACCTTGTCAAGGGATGAACCTCAAAAGTTCAAGATGAGCTGCTG 172
 QY 138 TCGGTGCGCGTCCGCGAGGGCTAAGGGGGCATCCCGGGGGCGGCTGCTCTCATGAC 197
 Db 173 ACAGTGAACCTGCGAAGGGCTATGGGGCATCCACGGGGGCGCTGCTGCAAGATGAC 232
 QY 198 GCGTTGACCTCAACGCAAGCTGATGCTTCTACCTGAGACCTTACGGGGCGAGCTC 257
 Db 233 GCCATATATCTCACTGACAACTTGTACGTAATCTGTAGTGTATGCTTGGAGCTC 292
 QY 258 ACCGCTAAGTGTCTGCGGACATGGGCTGAGAGATGGCGGGGAGCTGACAGCGGC 317
 Db 293 ACATGACTGTGTTAAGACATGGGCTTACAGAGGCTGGCTGACAGCTCAAAAG-- 349
 QY 318 ACGCAACAGGCTTGAAGCGCGCCAGCTGGAGATCCAGAGCCCTCTCTCATGCGAGCC 377
 Db 350 ACTAAAGAGAGTGTGAGGCTGTGAGCTGACAGCTGACAGCTGCTCAAGTACAGCC 409
 QY 378 AAGCCAGGCTGCACTTATATAGACCAAGACCGGGCTGGCTTATCGAGGGTCAAAAC 437
 Db 410 AGAACAGG---ACACTTTGTGACCAAGACAGCAAGCACTATGCGAGGTTACAGAA 466
 QY 438 GTTGAAGTGTCTGTGATGCTCTGTACGGGAAGTCTGACGAGTGAAGCAATACAGCA 497
 Db 467 GTGACGAGAGTGTGATGCTTTCATGAGGAGTGTGCTGACTGAAGACAGTACAGCA 526
 QY 498 GTGGGGGCGGAGCCCAACCAACCAAGCAAGATGGGAAGCTTTCAGTTTACACAGCC 557
 Db 527 GTTGTGTCAGAGACACACAGCAAGCAAGATGAGGAAGCTTTCAGTTTGTTCATCC 586
 QY 558 TGAACCTGACCTGCAAGAACTTGTCTCCCGAGGCCCTTAAGAGAGTCCAGTCCATCTG 617
 Db 587 TGAACCTGACCTGCAAGAACTTGTCTCCCGAGGCCCTTGAAGAAATACATCCCTTCTG 646
 QY 618 GTGAGGAGCTGAGCGAGAGCTGAGGCTCTT 649
 Db 647 GTGATGAGCTGAGAGAGAGTGAAGTATCTT 678

RESULT 4
 US-09-340-620A-62
 ; Sequence 62, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340, 620A
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US 09/245,281
 ; PRIOR FILING DATE: 1999-02-05
 ; PRIOR APPLICATION NUMBER: US 09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 62
 ; LENGTH: 579
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-340-620A-62

Query Match 42.7%; Score 316.2; DB 4; Length 579;
 Best Local Similarity 73.7%; Pred. No. 1.9e-62;
 Matches 431; Conservative 0; Mismatches 148; Indels 6; Gaps 2;
 Db 54 ATGGGGCGGGCGAGCGCCATCTTGAGTGGCTGAGAACTGACCGCGAGAGAGCTC 113
 1 ATGGGGCGGGCGAGAGATGCGATCTGAGCGCTCTTGAACCTTGTCAAGGGATGAACTC 60

QY 114 AAGAGTTCAAGCTGAGCTGTGTGCGTCCGCGCGAGGGCTACGGGGCATCCCG 173
 Db 61 AAAAGTTCAAGATGAAGCTGTGACATGTGCAATGCGAAGAGCTTATGGGGCATCCCA 120
 QY 174 CCGGGCGCGTGTCTGTCCATGAGAGCGCTTGGACCTTCAACGCAAGCTGTGAGCTTTCAC 233
 Db 121 CGGGGGCGCTGTGAGATGAGAGCGCATATGATCTGACCTGACAACTTGTCACTACTAT 180
 QY 234 CTGAGACCTTACGGGGCGGAGCTACCGCTTACCTGTCTGCGGACATGGGCTTCAAGAG 293
 Db 181 CTGAGATGATATGCTTGGAGCTTCAATGACTGTGCTTGAAGACATGGGCTTCAAGAG 240
 QY 294 ATGGCGGGAGCTGAGGGGGGCGGACGACGAGGGCTTGGAGCGGGCGAGCTGGAGATC 353
 Db 241 CTGCTGAGAGCTTGAAGG---ACTAAAGAGAGTGTGAGAGCTGTGAGCTTGAAGC 297
 QY 354 CAGGCGCTCTCTCACTGTCGAGCGGAGCGGAGCGGCTGCACTTATATAGACAGAGCGGCT 413
 Db 298 AGTGTCCCTGCTCAGAGTACAGGCAAGAG---ACACTTGTGACAGGACAGGCA 354
 QY 414 GCGCTTATCCGAGAGGTCAAAAGCTTGAAGTGTGCTGTGATGCTGTGACGGAAGTTC 473
 Db 355 GCACTCATTCGACGGGTCAAGAGTGAAGTGAAGTGTGATGCTTGTGATGAGCAAGTGTG 414
 QY 474 CTGACGATGAGAGTACAGGAGGAGGCGGCGGCGGCGGACCAACCAAGCAAGATGCGG 533
 Db 415 CTGACTAAGAGACAGTACAGGAGGCTGTGACAGAGACACACAGCAAGCAAGATGAGG 474
 QY 534 AAGCTTTCAGTTTACACAGGAGCTGAGACTGACCTGCAAGAGCTTCTCTCAGGCG 593
 Db 475 AAGCTTTCAGCTTGTTCATCTGAGACCTGACCTGCAAGAGCTCCCTCTCAGGCG 534
 QY 594 CTAAAGAGTCCAGTCTTACCTGTGTGAGAGAGCTTGAAGCGGAGC 638
 Db 535 TTGAAGAAATACATCCCTTCTGTGTGATGAGACTGAGAGAGAGC 579

RESULT 5
 US-09-252-991A-2958
 ; Sequence 2958, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 2958
 ; LENGTH: 909
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-2958

Query Match 7.2%; Score 53; DB 4; Length 909;
 Best Local Similarity 49.1%; Pred. No. 0.0027;
 Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
 QY 63 GCGGCGAGCGCCATCTGATGCGCTGAGAACTGACCGCGGAGAGCTCAAGAAAGTTC 122
 Db 576 GCGGCGCTCTCAACCAAGATACGAAGCGGCGAAGCGGCTCAAGGAAGTCCGCGCGTTC 635
 QY 123 AAGTGAAGCTGTCTGTGCGCGCTGCGGAGGAGCTTACGGGCGCATCCCGGGGGGCG 182
 Db 636 CGCTCTTACTGTGTGTGAGAGAGGTTTCAACACATGATGAGCTTCTCAGAGGCAAC 695
 QY 183 CTGTGTTCATGAGCGCTTGAAGCTTCAACGACAGAGCTGTGACGCTTCACTTGAAGACC 242

Db 696 GACCGCTACAGAACGCGCTGAGACAGCGCCGCAACCCGAGGCGCTTCATGACGAACTG 755
Qy 243 TACGGCGCCGAGCTCAACCGCTAACGTCGTGCGGACATGGGCGCTGACAGAGATGCGCGG 302
Db 756 CAGCGCGCGCGCTACACGACCGCATCCGAGTACGCCCGCAAGGTGGCGGACATGCGCCAGA 815
Qy 303 CAGCTGACGCGCGGACGACGACGAGGCTGTGAGCGCGCGCACT 347
Db 816 CAGATGACAGCTTACAGAGCGCTGCGCGCGCGCGGACGCGCGCT 860

RESULT 6

US-09-252-991A-3146
; Sequence 3146, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3146
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3146

Query Match 7.2%; Score 53; DB 4; Length 1218;
Best Local Similarity 49.1%; Pred. No. 0.0029;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
Qy 63 GCGCGGACGCCATCTGATGGCGTGAAGAACTGACCGCGGAGAGCTCAAGAACTTC 122
Db 925 GCGCGGCTCTCAACCGAATGAAAGGCGGCAAGCGGTCAAGAGAACTGCGCGCTTC 984
Qy 123 AAGCTGAAGCTGTGTCGTGTCGCGCTGCGGAGGCTTACCGGCGGCGCG 182
Db 985 CGCTCTACTGCTGCTTCGAGAGAGTTCACAGCATCTGCTTCACAGGCGAAC 1044
Qy 183 CTGCTGATGAGACCGCTTGAACCTTACCGACCAAGCTGTGAGCTTCTTACCTGAGACC 242
Db 1045 GACCGCTACCAAGCGCGCTGAGACGCGCGCAACCCGAGCGCTTCATGACGAACTG 1104
Qy 243 TACGGCGCCGAGCTCAACCGCTAACGTCGTGCGGACATGGGCGCTGACGAGATGCGCGG 302
Db 1105 CAGCGCGCGCGCTACACGCAACGATCCGAGTACGCCCGCAAGGTGGCGGACATGCGCCAGA 1164
Qy 303 CAGCTGACGCGCGGACGACGACGAGGCTGTGAGCGCGCGCACT 347
Db 1165 CAGATGACAGCTTACAGAGCGCTGCGCGCGCGGACGCGCGCT 1209

RESULT 7

US-09-252-991A-2822/c
; Sequence 2822, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2822
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2822

Query Match 7.2%; Score 53; DB 4; Length 1380;
Best Local Similarity 49.1%; Pred. No. 0.003;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGGACGCCATCTGATGGCGTGAAGAACTGACCGCGGAGAGCTCAAGAACTTC 122
Db 321 GCGCGGCTCTCAACCGAATGAAAGGCGGCAAGCGGTCAAGAGAACTGCGCGCTTC 262
Qy 123 AAGCTGAAGCTGTGTCGTGTCGCGCTGCGGAGGCTTACCGGCGGCGCG 182
Db 261 CGCTCTACTGCTGCTTCGAGAGAGTTCACAGCATCTGAGCTTCTTCACAGGCGAAC 202
Qy 183 CTGCTGATGAGACCGCTTGAACCTTACCGACCAAGCTGTGAGCTTCTTACCTGAGACC 242
Db 201 GACCGCTACCAAGCGCGCTGAGACGCGCGCAACCCGAGCGCTTCATGACGAACTG 142
Qy 243 TACGGCGCCGAGCTCAACCGCTAACGTCGTGCGGACATGGGCGCTGACGAGATGCGCGG 302
Db 141 CAGCGCGCGCGCTACCGCCACCGATCCGAGTACGCCCGCAAGGTGGCGGACATGCGCCAGA 82
Qy 303 CAGCTGACGCGCGGACGACGACGAGGCTGTGAGCGCGCGCACT 347
Db 81 CAGATGACAGCTTACAGAGCGCTGCGCGCGCGGACGCGCGCT 37

RESULT 8

US-09-252-991A-3050
; Sequence 3050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3050
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3050

Query Match 7.2%; Score 53; DB 4; Length 1500;
Best Local Similarity 49.1%; Pred. No. 0.003;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGGACGCCATCTGATGGCGTGAAGAACTGACCGCGGAGAGCTCAAGAACTTC 122
Db 920 GCGCGGCTCTCAACCGAATGAAAGGCGGCAAGCGGTCAAGAGAACTGCGCGCTTC 979
Qy 123 AAGCTGAAGCTGTGTCGTGTCGCGCTGCGGAGGCTTACCGGCGGCGCG 182
Db 980 CGCTCTACTGCTGCTTCGAGAGAGTTCACAGCATCTGAGCTTCTTCACAGGCGAAC 1039
Qy 183 CTGCTGATGAGACCGCTTGAACCTTACCGACCAAGCTGTGAGCTTCTTACCTGAGACC 242
Db 1040 GACCGCTACCAAGCGCGCTTGAACGCGCGCAACCCGAGCGCTTCATGACGAACTG 1099
Qy 243 TACGGCGCCGAGCTCAACCGCTAACGTCGTGCGGACATGGGCGCTGACGAGATGCGCGG 302
Db 1100 CAGCGCGCGCGCTACCGCCACCGATCCGAGTACGCCCGCAAGGTGGCGGACATGCGCCAGA 1159

Dy 303 CAGCTGCAGGGCGGCCACGACCAGGGCTCTGAAGCCGCCAGCT 347
 ||| | | | | | | | | | | | | | | |
Db 1160 CAGATGCAGACTTACCAAGCGGTGCGCGCGCGCACGCGCCT 1204

```

RESULT 9
US-09-252-991A-3097
: Sequence 3097, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 3097
: LENGTH: 2079
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-3097

```

	Query Match	5.8%;	Score 50.6;	DB 4;	Length 2079;
	Best Local Similarity	5.5%;	Pred. No. 0.011;		
	Matches 150;	Conservative	0;	Mismatches 144;	Indels 3; Gaps 1
QY	7	CGGCTGCACGCGGGGTGAGCGCGCGCAGCGCGCGCGGGAACTCTGAGACCAATGGGGCGCGCGC	66		
Db	401	CGGCGGAGCGCCCAATTTCAGGCGGATATGCGACCGCGGCTCAGCGGTAACTTGAGGGCGCGGC	460		
QY	67	GCGACGCCATTCCTGATATGCGCTGAGAACTGACCGCGCAGGAGCTCAAGAATTCAACG	126		
Db	461	GCCAGATCCACAGAGCGCGTTGCTGGAGGAGGCGAAGAACGATCCGCGCGCGCGCGCAGCC	520		
QY	127	TGAAGCTGCG---TGTTCGTTGCCGCTGCGCGAAGGGCTACGGGCGCATCCCGCGGGCGCGC	183		
Db	521	TGCTGCTGCGCTTCAACGAGTTGACAGCGCGGTGATGCTTACTAGACGCGGGTCCGCC	580		
QY	184	TGCTGTCCATGACGCGCTTGGACCTTCACCGACAAGCTGATCAGCTTCTTACCTGAGACCT	243		
Db	581	TGCTGTGTCAGATATTTCTCAAGACGCGCAGAGAACTGTGTTAACCCAGGGGGGCGAGGGCG	640		
QY	244	ACGGCGCGGAGCTCAACGCGCTTAACGTGCTTCGCGGAAATGGGCTGTGAGAGAGATGCGCG	300		
Db	641	ACGCGCTGAATTCGCTTCGCGCGGGTTCGGACACCTTCGCGCTTGCAGAGGTTCGCG	697		

```

RESULT 10
: US-09-144-085-3
: Sequence 3, Application US/09144085
: Patent No. 6280999
:
: GENERAL INFORMATION:
: APPLICANT: Gustafsson, Claes
: APPLICANT: Betlach, Mary C.
: APPLICANT: Ashley, Gary
: APPLICANT: Julian, Bryan
: APPLICANT: Ziemann, Rainer
: TITLE OF INVENTION: SORANOLIM POLYKETIDE SYNTHASES AND ENCODING DNA
: FILE REFERENCE: 30062-20020.20
: CURRENT APPLICATION NUMBER: US/09/144,085
: CURRENT FILING DATE: 1998-08-31
: EARLIER APPLICATION NUMBER: 09/010,809
: EARLIER FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 8
:
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 3

```

```

; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3

```

	Query Match	Best Local Similarity	51.6%;	Score 49.8;	DB 3;	Length 33529;	
	Matches	114;	Conservative	0;	Mismatches 107;	Indels 0;	Gaps 0;
QY	90	GAAACCTGACCGCCGAGAGCTCAAGAACTTCAAGCTGAAGCTGTGTCGTCGCCGCTG	149				
Db	14083	GAAACCGATGACCGGTGCGCCCGTCGCGACGTCGTGGAAGCCGAGAGAGGTCCCGCGTGTGCTG	14142				
QY	150	CGGAGAGGCTACCGGCGGCATCCCGCGCGGCGCGCGCTGTGCATGAGACGCTTGGACCTC	209				
Db	14143	GCCGGCATCGACCGCGCGCGCCCGCCCGCTGAGCGCGCGGTCTGACACTGGCCGCGCGCTCGAC	14202				
QY	210	ACCGACAAAGTGTGTCAGCTTCTTAAGCTGAGACCTTACGGCGCCGAGCTCACCGCTAACGTG	269				
Db	14203	GACGGCGTGTGTCGCGCGCCAGACGCGCGGACGCCCTCTCGCGGGTCTGGCGCCGAAGGTG	14262				
QY	270	CTGCGCGACATGGGCGCTTGAGAGAGATGGCGCGGACGCTGCA	310				
Db	14263	GACGGGCGCTGCACTTGTGACAGACTGACCGGAGACTGGA	14303				

```

RESULT 11
US-09-428-517-1/c
Sequence 1, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 50937
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-428-517-1

```

	Query Match	Best Local Similarity	6.6%; Score 49.2; DB 3; Length 50937;
	Matches 144; Conservative	0; Mismatches 158; Indels 0; Gaps 0;	47.7%; Pred. No. 0.047;
QY	86 GCTGGAGAACCTTGACCGCCGAGGAGCTCAAGAAATTCAAGCTGAAGTCTGTGCGTCC	145	
DB	1750 GCGGGAACCGCTGGCGGCGCAACCCGGCGTGCAGAGCGCTGATCCACTACCCGGGCGCGGTGCA	1691	
QY	146 GCTGCGGAGGAGGCTACCGGGCGCAATCCGCGGGGGCGGCTGTCTGCATGAGCGGCTTGGA	205	
DB	1690 CCGGTCCGAGGACGTTACCCCGGCAAGCCGAGGCGCGCACAGCCCGTGGCGGAGCG	1631	
QY	206 CCTTACCGACAAAGCTGTGTCACCTTCTTACCTTGAGAACTTACGCGCGCGAGCTCACCGTTAA	265	
DB	1630 GCTGCGCCCGGAGGAGTGTCTGAGCTGCCATAGGCGCGCACCCTGTCGACGACGCGCGTCAA	1571	
QY	266 CGTGCTCGCGGACATATGGGCTTCGAGAGATATGGCCGGGCACTGCAAGGCGGCGCAACGACA	325	
DB	1570 GCGCGTATGAGCGGATACGGGAGCGGCGGTGGCGCGCTCTGTGAATCCGGGCGCGCGCG	1511	

QY 326 GGGCTCGAGCGCGCCAGCTGGATCCAGGCCCCCTCTCACTGCGGACCCAGCCAGG 385
 Db 1510 TCGGCCCGGCGTATCCGCTGCTGGGTACGGGGCCGACTCTACGCGGGCCGGCGGC 1451
 QY 386 CC 387
 Db 1450 CC 1449

RESULT 12
 US-09-252-991A-1107/c
 ; Sequence 1107, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 1107
 ; LENGTH: 2610
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-1107

Query Match 6.6%; Score 49; DB 4; Length 2610;
 Best Local Similarity 49.8%; Pred. No. 0.027;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 62 CGCGCGGAGCGCCATCTGATGCGCTGAGAACTTACCGCCGAGAGCTCAAGAGTT 121
 Db 1942 CGAGCGGGGATTCGCTGTCAGCGCCCGGAGAGCTAACCTCCGCTGGCGAGCG 1883
 QY 122 CAAGCTGAAGCTGCTGCTGCTGCGCTGCGGAGGCTTACCGCGCGCGCGC 181
 Db 1882 CATCCGCGCACTGGCGCCGCGGACTTCTGTTCTCTTCTACTACGCGCGCTGCGCGC 1823
 QY 182 GCTGCTGCTCATGAGACCCCTTGAAGCTTACCGAGCAAGCTGTGAGCTTACTGAGAG 241
 Db 1822 CGAGCTGCTGCTGCTGCTGCGCGGAGGCTTACCACTTCAAGCTTCAAGCTGCTGCGCG 1763
 QY 242 CTACGCGCGCGAGCTCAACGCTTACGCTGCGGAGATGAGGCGCTGCGCGCG 301
 Db 1762 CTACCGCGGAGCGAGCGCCCGCGGAGCTGCTGCTGCTCAACGCGGAGAGCGAGG 1703
 QY 302 GCAGCTGCA 310
 Db 1702 GACCTGCA 1694

RESULT 13
 US-09-252-991A-977
 ; Sequence 977, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 977

; LENGTH: 3114
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-977

Query Match 6.6%; Score 49; DB 4; Length 3114;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 62 CGCGCGGAGCGCCATCTGATGCGCTGAGAACTTACCGCCGAGAGCTCAAGAGTT 121
 Db 1281 CGAGCGGGGATTCGCTGTCAGCGCCCGGAGAGCTAACCTCCGCTGGCTGAGCG 1340
 QY 122 CAAGCTGAAGCTGCTGCTGCTGCGCTGCGGAGGCTTACCGCGCGCGCGC 181
 Db 1341 CATCCGCGCACTGGCGCCGCGGACTTCTGTTCTCTTCTACTACGCGCGCTGCTGCGCGC 1400
 QY 182 GCTGCTGCTCATGAGACCCCTTGAAGCTTACCGAGCAAGCTGTGAGCTTACTGAGAG 241
 Db 1401 CGAGCTGCTGCTGCTGCGCGGAGGCTTACCACTTCAAGCTTCAAGCTGCTGCGCGC 1460
 QY 242 CTACGCGCGCGAGCTCAACGCTTACGCTGCGGAGATGAGGCGCTGCGCGCG 301
 Db 1461 CTACCGCGGAGCGAGCGCCCGCGGAGCTGCTGCTGCTCAACGCGGAGAGCGAGG 1520
 QY 302 GCAGCTGCA 310
 Db 1521 GACCTGCA 1529

RESULT 14
 US-09-252-991A-937
 ; Sequence 937, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 937
 ; LENGTH: 3195
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-937

Query Match 6.6%; Score 49; DB 4; Length 3195;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 62 CGCGCGGAGCGCCATCTGATGCGCTGAGAACTTACCGCCGAGAGCTCAAGAGTT 121
 Db 478 CGAGCGGGGATTCGCTGTCAGCGCCCGGAGAGCTAACCTCCGCTGGCTGAGCG 537
 QY 122 CAAGCTGAAGCTGCTGCTGCTGCGCTGCGGAGGCTTACCGCGCGCGCGC 181
 Db 538 CATCCGCGCACTGGCGCCGCGGACTTCTGTTCTCTTCTACTACGCGCGCTGCTGCGCGC 597
 QY 182 GCTGCTGCTCATGAGACCCCTTGAAGCTTACCGAGCAAGCTGTGAGCTTACTGAGAG 241
 Db 598 CGAGCTGCTGCTGCTGCGCGGAGGCTTACCACTTCAAGCTTCAAGCTGCTGCGCGC 657
 QY 242 CTACGCGCGCGAGCTCAACGCTTACGCTGCGGAGATGAGGCGCTGCGCGCG 301
 Db 658 CTACCGCGGAGCGAGCGCCCGCGGAGCTGCTGCTGCTCAACGCGGAGAGCGAGG 717
 QY 302 GCAGCTGCA 310

Db 718 GACCCTGCA 726

RESULT 15
US-09-252-991A-10961/c
; Sequence 10961, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10961
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10961

Query Match 6.6%; Score 48.8; DB 4; Length 468;
Best Local Similarity 47.4%; Pred. No. 0.02;
Matches 146; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGCGGGGTGAGCGGGCGACGCGCGGGGATCTGGAGCCATGGGGC 60
Db 325 CGCTGCCGGTGCACCGAGGTTTCGACGAGTTGAGGCCGCGGTGCTGAGAGCAGGTCC 266
QY 61 GCGGCGCGGACGCCATCTGTGATGCGCTGAGAACCTGACCGCCGAGGAGCTCAAGAACT 120
Db 265 GCGATGCCGCGCTGTTCCGCGCGCGCGCGCGCGCGCTGCTGACGACGACGAGCTTGC 206
QY 121 TCAGCTGAAGCTGTGCTGCGTGCCTGCGCGAGGAGGCTACGAGGCGCATCCGCGGGCG 180
Db 205 GCGAGCTCCCGTCCAGGCGACGCGCCCGACGAAGTCATCAGATTGCGCTGCCGAGCGCCC 146
QY 181 CGCTGCTGTCCATGAGAGCGCTTGGACCTCACCGACAGAGCTGCTTACCTGAGAGA 240
Db 145 CGGCGAGCTGCGCGCGCGCGCGCGCTGCTGACCTGCCCTGCTGACGACGAGTCTCCGG 86
QY 241 CCTAGCGCGCGAGCTCAACGCTACGCTGCGGACATGAGGCTCTGAGAGAGATGCGCG 300
Db 85 CCGGCGCGCGCGAGCGCTTCCAGTGCCTGCTGTTGACGTGCCCGGCTGACCTGCGCG 26
QY 301 GGCAGCTG 308
Db 25 TGCCGCTG 18

Search completed: January 29, 2004, 13:11:29
Job time : 41.6087 secs

Db	61	GGCGGCGCCACGGCACTCTTGATATGGCGCTGGAGAACCTTACCGCGGAGGACTCAAGAGT	120
Qy	121	TCAAAGCTGAAGCTGCTGTGGATGCCGCTGCGGAGGAGCTACGAGGCGCATCCCGGCGGCG	180
Db	121	TCAAAGCTGAAGCTGCTGTGGATGCCGCTGCGGAGGAGCTACGAGGCGCATCCCGGCGGCG	180
Qy	181	CGGTGCTGTTCATGAGACGCTTTGAGACTTCACCGACAACTGCTGACGTTTCACTTGGAGA	240
Db	181	CGGTGCTGTTCATGAGACGCTTTGAGACTTCACCGACAACTGCTGACGTTTCACTTGGAGA	240
Qy	241	CTTACGAGCGCCGAGCTCACCGGCTTAACTGTCGCGGACATAGGGGCTGAGAGATGGCGG	300
Db	241	CTTACGAGCGCCGAGCTCACCGGCTTAACTGTCGCGGACATAGGGGCTGAGAGATGGCGG	300
Qy	301	GGCAGCTGACAGGCGGCGCCACGACACAGAGGCTTGGAGCGCGCAGCTGAGATCCAGGCCC	360
Db	301	GGCAGCTGACAGGCGGCGCCACGACACAGAGGCTTGGAGCGCGCAGCTGAGATCCAGGCCC	360
Qy	361	CTTCTCAATGTCGGACAGCCAGCCAGGCTTGCACTTTATATGACACGACCCGGGCTGCTTA	420
Db	361	CTTCTCAATGTCGGACAGCCAGCCAGGCTTGCACTTTATATGACACGACCCGGGCTGCTTA	420
Qy	421	TGCGAGGGGTCAAAACGTTGATGGCTGCTGAGATGCTCTGTACGGGAAAGTCTCTGACGG	480
Db	421	TGCGAGGGGTCAAAACGTTGATGGCTGCTGAGATGCTCTGTACGGGAAAGTCTCTGACGG	480
Qy	481	ATGAGCAGTACACAGGCACTGCGGCGCGAGCCACCAACCCAAAGACAGTGGAGAGTCT	540
Db	481	ATGAGCAGTACACAGGCACTGCGGCGCGAGCCACCAACCCAAAGACAGTGGAGAGTCT	540
Qy	541	TCAGTTTCAACACCAAGCTTGGAACTGAGACTTGACAGGACTTCTCTCCAGGCGCTTAAAGG	600
Db	541	TCAGTTTCAACACCAAGCTTGGAACTGAGACTTGACAGGACTTCTCTCCAGGCGCTTAAAGG	600
Qy	601	AGTCCCAAGTCTTAACTGCTGGTGGAGACCTTGGAGCGGAGCTGAGGCTCTTCCACGACAC	660
Db	601	AGTCCCAAGTCTTAACTGCTGGTGGAGACCTTGGAGCGGAGCTGAGGCTCTTCCACGACAC	660
Qy	661	TCCGCTCAGCCCTCGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAAT	720
Db	661	TCCGCTCAGCCCTCGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAAT	720
Qy	721	ATACGAAAAAGCCAGCTTGA 740	
Db	721	ATACGAAAAAGCCAGCTTGA 740	
RESULT 2			
US-09-996-617-7			
Sequence 7, Application US/09996617			
Patent No. US20020128198A1			
GENERAL INFORMATION:			
APPLICANT: Bertin, John			
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED			
FILE REFERENCE: 07334-340001			
CURRENT APPLICATION NUMBER: US/09/996,617			
CURRENT FILING DATE: 2001-11-27			
PRIORITY APPLICATION NUMBER: 09/931,071			
PRIORITY FILING DATE: 2001-08-15			
PRIORITY APPLICATION NUMBER: 09/428,252			
PRIORITY FILING DATE: 1999-10-27			
PRIORITY APPLICATION NUMBER: 09/340,620			
PRIORITY FILING DATE: 1999-06-28			
NUMBER OF SEQ. ID NOS: 10			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ. ID NO 7			
LENGTH: 740			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			

LOCATION: (54)... (638
US-09-996-617-7

Query Match	100.0%	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0%	Pred. No. 1.1e-196;		
Matches 740; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0

QY	1	CGGCTCGGCTCAGCGGGGTATGCGCGGAGCGCCGCGGGATCTCTGAGCACTAGGAGC	60
Db	1	CGGCTCGGCTCAGCGGGGTATGCGCGGAGCGCCGCGGGATCTCTGAGCACTAGGAGC	60
QY	61	GGCGCGCGACGCCATCTCTGATGCGCTTGAGAACTTGACCGCCGAGAGCTCAAGAGT	120
Db	61	GGCGCGCGACGCCATCTCTGATGCGCTTGAGAACTTGACCGCCGAGAGCTCAAGAGT	120
QY	121	TCAGGTGAAGGTGCTGTGCGGTGCGCGCTGCGCGAGGGCTACGGGGGCAATCCCGGGGCG	180
Db	121	TCAGGTGAAGGTGCTGTGCGGTGCGCGCTGCGCGAGGGCTACGGGGGCAATCCCGGGGCG	180
QY	181	CGGTGCTGTCAATGACCGCTTGACCTTCAACGAACTGCTGAGCTTCACTTGAGA	240
Db	181	CGGTGCTGTCAATGACCGCTTGACCTTCAACGAACTGCTGAGCTTCACTTGAGA	240
QY	241	CCTACGGCGCCGAGCTCAACCGCTTAACGTGTCTGCGGACATGGGCTTGAGAGATGGCCG	300
Db	241	CCTACGGCGCCGAGCTCAACCGCTTAACGTGTCTGCGGACATGGGCTTGAGAGATGGCCG	300
QY	301	GGCAGGTGAGGGCGGCGCAACGACCAAGGGCTTGAGCGCGCCAGCTGGATTCAGGCC	360
Db	301	GGCAGGTGAGGGCGGCGCAACGACCAAGGGCTTGAGCGCGCCAGCTGGATTCAGGCC	360
QY	361	CTCTCAGTGCGGACGCCAAGCCAGGCTTGCACTTTATAGACAGACCCGGCTGCGCTTA	420
Db	361	CTCTCAGTGCGGACGCCAAGCCAGGCTTGCACTTTATAGACAGACCCGGCTGCGCTTA	420
QY	421	TGCGCAGGGGTCACAAAGTTGATGCTGGTCTGAAATGCTGTAGAGGGAAGGTCTGACGG	480
Db	421	TGCGCAGGGGTCACAAAGTTGATGCTGGTCTGAAATGCTGTAGAGGGAAGGTCTGACGG	480
QY	481	ATGAGCAGTACAGGACAGTGGCGGGCCGAGCCACCAACCCAAAGCAATGTCGAAAGCTT	540
Db	481	ATGAGCAGTACAGGACAGTGGCGGGCCGAGCCACCAACCCAAAGCAATGTCGAAAGCTT	540
QY	541	TGAGTTTCAACCAACCTTGGAATCTGGAACCTGCAAGAACTTGTCTCTCAAGCCTTAAGGG	600
Db	541	TGAGTTTCAACCAACCTTGGAATCTGGAACCTGCAAGAACTTGTCTCTCAAGCCTTAAGGG	600
QY	601	AGTCCAGTCTACCTGTGTGAGAGACTGGAAGCGGAGCTGAGGCTCTTCCACAGCAAC	660
Db	601	AGTCCAGTCTACCTGTGTGAGAGACTGGAAGCGGAGCTGAGGCTCTTCCACAGCAAC	660
QY	661	TCCGCTCAGCCCTTGGAATCTCCACCAATCATCTGTAATGTATCTTTTATACAAAT	720
Db	661	TCCGCTCAGCCCTTGGAATCTCCACCAATCATCTGTAATGTATCTTTTATACAAAT	720
QY	721	ATACGAAAAGCCAGCTTGA 740	
Db	721	ATACGAAAAGCCAGCTTGA 740	

RESULT 3
 US-09-996-617-10/c
 ; Sequence 10, Application US/09996617
 ; Patent No. US20020128198A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-34001
 ; CURRENT APPLICATION NUMBER: US/09/996,617
 ; PRIOR FILING DATE: 2001-11-27
 ; PRIOR APPLICATION NUMBER: 09/931,071
 ; PRIOR FILING DATE: 2001-08-15

```

: PRIOR APPLICATION NUMBER: 09/428,252
: PRIOR FILING DATE: 1999-10-27
: PRIOR APPLICATION NUMBER: 09/340,620
: PRIOR FILING DATE: 1999-06-28
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 10
: LENGTH: 740
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-996-617-10

```

Query Match	100.0%;	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0%;	Pred. No. 1.1e-196;		
Matches 740;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	CGGCTCCGGCTGCAAGCGGGGTATAGAGCGCGGACAGCGCGCGGGAAATCTTGAGACCATGGGCG	60
Db	740	CGGCTCCGGCTGCAAGCGGGGTATAGAGCGCGGACAGCGCGCGGGAAATCTTGAGACCATGGGCG	681
QY	61	GGCGGCGCAGCAGCCATCTTGATGGCGTTGGAGAACTTGACCCGACGAGAGCTCAAGAGT	120
Db	680	GGCGGCGCAGCAGCCATCTTGATGGCGTTGGAGAACTTGACCCGACGAGAGCTCAAGAGT	621
QY	121	TCAAGCTGAAGCTGCTGTGTGGTCCGCTGGCGAGAGGCTTACGGGCGCATCCCGCGGGCG	180
Db	620	TCAAGCTGAAGGCTGCTGTGTGGTCCGCTGGCGAGAGGCTTACGGGCGCATCCCGCGGGCG	561
QY	181	CGGTGCTGTCCATGAGACCGCTTGGACCTCACCGAACAAGCTGTAGCTTTTACTGTGAGA	240
Db	560	CGGTGCTGTCCATGAGACCGCTTGGACCTCACCGAACAAGCTGTAGCTTTTACTGTGAGA	501
QY	241	CCATCCGGCGCGCAGAGCTCACCGCTTAACTGCTGGCGGACATGGGCTTGACGAGATGGCG	300
Db	500	CCATCCGGCGCGCAGAGCTCACCGCTTAACTGCTGGCGGACATGGGCTTGACGAGATGGCG	441
QY	301	GGCAGCTGCAAGGCGGCGACGACACCAAGGGCTTGAGAGCGCGCCACGCTGGATTCAGGCC	360
Db	440	GGCAGCTGCAAGGCGGCGACGACACCAAGGGCTTGAGAGCGCGCCACGCTGGATTCAGGCC	381
QY	361	CTGCTCAAGTCGGGAGCGCAAGCGAGGCTGCGACTTTATATGACACAGCACCGGCTGGCTTA	420
Db	380	CTGCTCAAGTCGGGAGCGCAAGCGAGGCTGCGACTTTATATGACACAGCACCGGCTGGCTTA	321
QY	421	TGCGAGGGGTCACAAACTTTGATGGCTGCTGATGCTCTGTACGGGAAAGTCTTGACGG	480
Db	320	TGCGAGGGGTCACAAACTTTGATGGCTGCTGATGCTCTGTACGGGAAAGTCTTGACGG	261
QY	481	ATGAGCAGTACCAAGGCAGTGCGGGCTCGAGCCCAACCCAAAGCAAGATGGAGGCTCT	540
Db	260	ATGAGCAGTACCAAGGCAGTGCGGGCTCGAGCCCAACCCAAAGCAAGATGGAGGCTCT	201
QY	541	TCAAGTTTCAACCAAGCTGGAGACTGAGACCTGCAAGAGATTTGCTCTCTCAAGGCTTAAAGG	600
Db	200	TCAAGTTTCAACCAAGCTGGAGACTGAGACCTGAGCACTTCTCTCTCAAGGCTTAAAGG	141
QY	601	AGTCCAGTCTACCTCGGTGAGAGACTGAGACGAGCTGAGGCTCTCTCCACAGCAAC	660
Db	140	AGTCCAGTCTACCTCGGTGAGAGACTGAGACGAGCTGAGGCTCTCTCCACAGCAAC	81
QY	661	TCCGATCAGCCCTGGCAATCCACCAATCATCTGTATCTGATCTTTTATACAAAT	720
Db	80	TCCGATCAGCCCTGGCAATCCACCAATCATCTGTATCTGATCTTTTATACAAAT	21
QY	721	ATACGAAAAAGCAGCTTGA	740
Db	20	ATACGAAAAAGCAGCTTGA	1

```

1  GENERAL INFORMATION:
2  APPLICANT: Bertin, John
3  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
4  FILE REFERENCE: 07334-330001
5  CURRENT APPLICATION NUMBER: US/09/841,879B
6  CURRENT FILING DATE: 2001-04-24
7  PRIOR APPLICATION NUMBER: US 09/728,721
8  PRIOR FILING DATE: 2000-12-01
9  PRIOR APPLICATION NUMBER: US 09/340,620
10 PRIOR FILING DATE: 1999-06-28
11 NUMBER OF SEQ ID NOS: 19
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 4
14 LENGTH: 740
15 TYPE: DNA
16 ORGANISM: Homo sapiens
17 FEATURE:
18 NAME/KEY: CDS
19 LOCATION: (54)...(638)
20 US-03-841-879B-4

```

Query Match	100.0%;	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0%;	Pred. No. 1.1e-196;		
Matches 740;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	CGGCTCGGCTGCACCGGGGGTGAAGGCGGCGAGCGCCGGGAATCTTGAGCAATGGAGC	60
Db	1	CGGCTCGGCTGCACCGGGGGTGAAGGCGGCGAGCGCCGGGAATCTTGAGCAATGGAGC	60
Qy	61	GCGCGCACGACCCATCTGTGAATGCGCTTGAAGAACTTGAACCGCGAGAGCTCAAGAGT	120
Db	61	GCGCGCACGACCCATCTGTGAATGCGCTTGAAGAACTTGAACCGCGAGAGCTCAAGAGT	120
Qy	121	TCAGGTGAAGCTGCTGTGCGTCCGCTGCGCGAGAGGGCTACGGGGCGATCCCGGGGGCG	180
Db	121	TCAGGTGAAGCTGCTGTGCGTCCGCTGCGCGAGAGGGCTACGGGGCGATCCCGGGGGCG	180
Qy	181	CGTGTGTCACATGGAACGCTTGAACCTTCAACGGAACAAGTGGTCAAGCTTCACTTGAGA	240
Db	181	CGTGTGTCACATGGAACGCTTGAACCTTCAACGGAACAAGTGGTCAAGCTTCACTTGAGA	240
Qy	241	CTTACGGGCGCCGAGCTCACCGCTTAACTGTGTGCGCGACATGAGGCTGTCAAGAAATGGCCG	300
Db	241	CTTACGGGCGCCGAGCTCACCGCTTAACTGTGTGCGCGACATGAGGCTGTCAAGAAATGGCCG	300
Qy	301	GGCAGCTGAGAGCGGCGCACGACCAAGGGCTTGGAGCGCGCCAGCTGGGATCCAGGCC	360
Db	301	GGCAGCTGAGAGCGGCGCACGACCAAGGGCTTGGAGCGCGCCAGCTGGGATCCAGGCC	360
Qy	361	CTTCTCAGTCGGCAGCCCAAGCCAGGCTTGCATTTATATGACACAGCACCGGGCTGCGCTTA	420
Db	361	CTTCTCAGTCGGCAGCCCAAGCCAGGCTTGCATTTATATGACACAGCACCGGGCTGCGCTTA	420
Qy	421	TTCGCGAGGGTTCACAAACGTGATGAGGGCTGAGATGCTGTGTAAGGGAAAGTCTTGACGG	480
Db	421	TTCGCGAGGGTTCACAAACGTGATGAGGGCTGAGATGCTGTGTAAGGGAAAGTCTTGACGG	480
Qy	481	ATGAGCAGTACCAAGCAGATGCGGGCCGAGCCCAACCAACGCAAGATGCGAAGCTCT	540
Db	481	ATGAGCAGTACCAAGCAGATGCGGGCCGAGCCCAACCAACGCAAGATGCGAAGCTCT	540
Qy	541	TCAGTTTCACACCAAGCTTGAACTGGAACCTTCAAGAGCTTCTTCAAGGCTTAAAGG	600
Db	541	TCAGTTTCACACCAAGCTTGAACTGGAACCTTCAAGAGCTTCTTCAAGGCTTAAAGG	600
Qy	601	AGTCCAGATCTTACCTGTGTGAGACCTGAGACGGAAGCTGAGGCTCTTCCACAGCAAC	660
Db	601	AGTCCAGATCTTACCTGTGTGAGACCTGAGACGGAAGCTGAGGCTCTTCCACAGCAAC	660
Qy	661	TTCGGTCAAGCCCTTGCAATCCCAACCAATATCTCTGAATCTTTTATATACAAAT	720
Db	661	TTCGGTCAAGCCCTTGCAATCCCAACCAATATCTCTGAATCTTTTATATACAAAT	720

QY 421 TCCGAGAGGTCACAAGCTTGAAGTGGCTGTGATGCTCTGTATAGGGAAGGCTCTGACGG 480
DB 421 TCCGAGAGGTCACAAGCTTGAAGTGGCTGTGATGCTCTGTATAGGGAAGGCTCTGACGG 480
QY 481 ATGAGCAGTACCAAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGGAAGCTCT 540
DB 481 ATGAGCAGTACCAAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGGAAGCTCT 540
QY 541 TCACTTTCACACCAAGCTGTGAAGTGAAGCTGTGCAAGAGCTTGTCTCCAGGCTTAAAGG 600
DB 541 TCACTTTCACACCAAGCTGTGAAGTGAAGCTGTGCAAGAGCTTGTCTCCAGGCTTAAAGG 600
QY 601 AGTCCAGCTCTACCTGTGTGAAGAGCTGTGAAGCTGTGAGCTGTCTCCAGCAACAC 660
DB 601 AGTCCAGCTCTACCTGTGTGAAGAGCTGTGAAGCTGTGAGCTGTCTCCAGCAACAC 660
QY 661 TCCGTCAGGCTCTGAGCAATCCCAACCAATCATCTGATCTTTTATACAAAT 720
DB 661 TCCGTCAGGCTCTGAGCAATCCCAACCAATCATCTGATCTTTTATACAAAT 720
QY 721 ATACGAAAGCCAGCTTGAA 740
DB 721 ATACGAAAGCCAGCTTGAA 740

RESULT 7

US-10-240-145-12
; Sequence 12, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 12
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(632)
US-10-240-145-12

Query Match 99.1%; Score 733; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 9.7e-195;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GGCTGCAGCGGGGTGAGCGCGGCGAGCGCGCGGGATCTTGAGCCATGCGGCGCGCGCG 67
DB 2 GGCTGCAGCGGGGTGAGCGCGGCGAGCGCGCGGGATCTTGAGCCATGCGGCGCGCGCG 61
QY 68 CGAGCCATCTGATGATGCGCTGGAAGCTGACCGCGGAGAGAGTCAAGAGTCAAGCT 127
DB 62 CGAGCCATCTGATGATGCGCTGGAAGCTGACCGCGGAGAGTCAAGAGTCAAGCT 121
QY 128 GAAGCTCTGTGCTGCTGCTGCGGAGGAGCTACGCGGCGCATCCCGCGGCGCGCTGCT 187

DB 122 GAAGCTCTGTGCTGCTGCTGCGGAGGAGCTACGCGGCGCATCCCGCGGCGCGCTGCT 181
QY 188 GTCCATGAGAGCCTTGTGAGCTTCAACCGAAGCTGTGAGCTTCTTCACTGAGAGCTTACG 247
DB 182 GTCCATGAGAGCCTTGTGAGCTTCAACCGAAGCTGTGAGCTTCTTCACTGAGAGCTTACG 241
QY 248 GCGGAGCTACCGCTTGAAGTGTGCGGAGCATGAGGCTGAGAGATGCGCGGAGCT 307
DB 242 GCGGAGCTACCGCTTGAAGTGTGCGGAGCATGAGGCTGAGAGATGCGCGGAGCT 301
QY 308 GCAAGCGGAGCAGCAGCAGGAGCTTGTGAGCGGCGGAGCTGAGATCCAGGCGCTCTCA 367
DB 302 GCAAGCGGAGCAGCAGCAGGAGCTTGTGAGCGGCGGAGCTGAGATCCAGGCGCTCTCA 361
QY 368 GTGCGAGCCAGCAGCAGCTTGAAGCTTATAGACAGCAGCAGGCTGCTTATGCGAG 427
DB 362 GTGCGAGCCAGCAGCAGCTTGAAGCTTATAGACAGCAGCAGGCTGCTTATGCGAG 421
QY 428 GGTCAACAAGTGTGAGTGTGCTGAGATGCTGTATAGGGAAGTCTCTGACGAGTGAACA 487
DB 422 GGTCAACAAGTGTGAGTGTGCTGAGATGCTGTATAGGGAAGTCTCTGACGAGTGAACA 481
QY 488 GTACAGAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGGAAGCTCTTCAAGTT 547
DB 482 GTACAGAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGGAAGCTCTTCAAGTT 541
QY 548 CACACAGCCTGGAAGCTGACCTTGAAGAGCTTGTCTTCCAGGCTTAAAGAGTCCCA 607
DB 542 CACACAGCCTGGAAGCTGACCTTGAAGAGCTTGTCTTCCAGGCTTAAAGAGTCCCA 601
QY 608 GTCTTACCTGTGAGAGAGCTGTGAGCGGAGCTGAGCTCTTCCAGCAACACTCCGCT 667
DB 602 GTCTTACCTGTGAGAGAGCTGTGAGCGGAGCTGAGCTCTTCCAGCAACACTCCGCT 661
QY 668 AGCCCTGCAATCCCAACCAATCATCTGATCTTTTATACAAATTAAGAA 727
DB 662 AGCCCTGCAATCCCAACCAATCATCTGATCTTTTATACAAATTAAGAA 721
QY 728 AAGCCAGCTTGAA 740
DB 722 AAGCCAGCTTGAA 734

RESULT 8

US-10-131-410-21
; Sequence 21, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-21

Query Match 99.1%; Score 733; DB 12; Length 779;
Best Local Similarity 100.0%; Pred. No. 9.8e-195;

Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 7 CGGCTGAGCGGGGTGAGCGCGGAGCGGCGGGATCTTGAGCCATGCGCGC 66
    |||
Db 47 CGGCTGAGCGGGGTGAGCGCGGAGCGGCGGGATCTTGAGCCATGCGCGC 106
    |||
QY 67 GCACCGCATCTCTGAGCGCTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 126
    |||
Db 107 GCACCGCATCTCTGAGCGCTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 166
    |||
QY 127 TGAAGCTGTGTGCGTCCGCTGCGGAGGAGCTTACCGCGGAGGCGCGTGC 186
    |||
Db 167 TGAAGCTGTGTGCGTCCGCTGCGGAGGAGCTTACCGCGGAGGCGCGTGC 226
    |||
QY 187 TGTCCATGAGACGCTTGAGCTTACCGGACAGCTTGACCTTCTTACCTGAGAC 246
    |||
Db 227 TGTCCATGAGACGCTTGAGCTTACCGGACAGCTTGACCTTCTTACCTGAGAC 286
    |||
QY 247 GCGCCGAGCTCACCGCTTAAGTGTGCGGACATGCGGCTTGAGAGATGCGCGGAGC 306
    |||
Db 287 GCGCCGAGCTCACCGCTTAAGTGTGCGGACATGCGGCTTGAGAGATGCGCGGAGC 346
    |||
QY 307 TGCAGGCGGCGACGACGACGAGGCTCTGAGAGCGGCGGAGTCAAGCCCTCTC 366
    |||
Db 347 TGCAGGCGGCGACGACGACGAGGCTCTGAGAGCGGCGGAGTCAAGCCCTCTC 406
    |||
QY 367 AGTGGAGCGCAAGCGGCTTGACCTTTATAGCCAGCAACCGGCTGCGCTTATCGCA 426
    |||
Db 407 AGTGGAGCGCAAGCGGCTTGACCTTTATAGCCAGCAACCGGCTGCGCTTATCGCA 466
    |||
QY 427 GGGTCACAAACGTTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATG 486
    |||
Db 467 GGGTCACAAACGTTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATG 526
    |||
QY 487 AGTACCGAGGAGTGCAGGCGGAGCGGCGGACCAACCAAGAGATGCGGAGCTTACGTT 546
    |||
Db 527 AGTACCGAGGAGTGCAGGCGGAGCGGCGGACCAACCAAGAGATGCGGAGCTTACGTT 586
    |||
QY 547 TCACACGAGCTTGAGACTGAGACTTGACCTTCTTCAAGGCTTAAAGGAGTCCC 606
    |||
Db 587 TCACACGAGCTTGAGACTGAGACTTGACCTTCTTCAAGGCTTAAAGGAGTCCC 646
    |||
QY 607 AGTCTTACCTGTGAGAGGAGCTTGAGAGGAGCTTGAGAGGCTTGAGAGGCTTGAGAGG 666
    |||
Db 647 AGTCTTACCTGTGAGAGGAGCTTGAGAGGAGCTTGAGAGGCTTGAGAGGCTTGAGAGG 706
    |||
QY 667 CAGCCCTGCGCAATCCACCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
    |||
Db 707 CAGCCCTGCGCAATCCACCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 766
    |||
QY 727 AAAGCCAGCTTGA 739
    |||
Db 767 AAAGCCAGCTTGA 779
    |||

```

RESULT 9
US-10-106-698-1144/c

```

; Sequence 1144, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0

```

; SEQ ID NO 1144
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-1144

Query Match 97.5%; Score 721.2; DB 15; Length 806;
Best Local Similarity 99.6%; Pred. No. 2e-191;
Matches 731; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

```

QY 7 CGGCTGAGCGGGGTGAGCGCGGAGCGGCGGGATCTTGAGCCATGCGCGC 66
    |||
Db 739 CGGCTGAGCGGGGTGAGCGCGGAGCGGCGGGATCTTGAGCCATGCGCGC 680
    |||
QY 67 GCACCGCATCTCTGAGCGCTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 126
    |||
Db 679 GCACCGCATCTCTGAGCGCTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 620
    |||
QY 127 TGAAGCTGTGTGCGTCCGCTGCGGAGGAGCTTACCGCGGAGGCGCGTGC 186
    |||
Db 619 TGAAGCTGTGTGCGTCCGCTGCGGAGGAGCTTACCGCGGAGGCGCGTGC 560
    |||
QY 187 TGTCCATGAGACGCTTGAGCTTACCGGACAGCTTGACCTTCTTACCTGAGAC 246
    |||
Db 559 TGTCCATGAGACGCTTGAGCTTACCGGACAGCTTGACCTTCTTACCTGAGAC 500
    |||
QY 247 GCGCCGAGCTCACCGCTTAAGTGTGCGGACATGCGGCTTGAGAGATGCGCGGAGC 306
    |||
Db 499 GCGCCGAGCTCACCGCTTAAGTGTGCGGACATGCGGCTTGAGAGATGCGCGGAGC 440
    |||
QY 307 TGCAGGCGGCGACGACGACGAGGCTCTGAGAGCGGCGGAGTCAAGCCCTCTC 366
    |||
Db 439 TGCAGGCGGCGACGACGACGAGGCTCTGAGAGCGGCGGAGTCAAGCCCTCTC 380
    |||
QY 367 AGTGGAGCGCAAGCGGCTTGACCTTTATAGCCAGCAACCGGCTGCGCTTATCGCA 426
    |||
Db 379 AGTGGAGCGCAAGCGGCTTGACCTTTATAGCCAGCAACCGGCTGCGCTTATCGCA 320
    |||
QY 427 GGGTCACAAACGTTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATG 486
    |||
Db 319 GGGTCACAAACGTTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATG 260
    |||
QY 487 AGTACCGAGGAGTGCAGGCGGAGCGGCGGACCAACCAAGAGATGCGGAGCTTACGTT 546
    |||
Db 259 AGTACCGAGGAGTGCAGGCGGAGCGGCGGACCAACCAAGAGATGCGGAGCTTACGTT 200
    |||
QY 547 TCACACGAGCTTGAGACTGAGACTTGACCTTCTTCAAGGCTTAAAGGAGTCCC 606
    |||
Db 199 TCACACGAGCTTGAGACTGAGACTTGACCTTCTTCAAGGCTTAAAGGAGTCCC 140
    |||
QY 607 AGTCTTACCTGTGAGAGGAGCTTGAGAGGAGCTTGAGAGGCTTGAGAGGCTTGAGAGG 666
    |||
Db 139 AGTCTTACCTGTGAGAGGAGCTTGAGAGGAGCTTGAGAGGCTTGAGAGGCTTGAGAGG 80
    |||
QY 667 CAGCCCTGCGCAATCCACCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
    |||
Db 79 CA-CCCTGCGCAATCCACCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 21
    |||
QY 727 AAAGCCAGCTTGA 740
    |||
Db 20 AAAGCCAGCTTGA 7
    |||

```

RESULT 10
US-09-925-301-278

```

; Sequence 278, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 9.48584 Seconds

(without alignments)
3262.936 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARALDALNTREEL.....LLQLARESGSYLVIEDIERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	990	100.0	195	20 AA198553 Human breast tumour
2	990	100.0	195	22 AA068525 Human novel cytochrome
3	990	100.0	195	22 AA060588 Human target of me
4	990	100.0	195	22 AA020085 Human CARD-5 prote
5	990	100.0	195	23 AA093353 Human caspase recr
6	990	100.0	195	23 AA017854 Pyrin domain conta
7	990	100.0	195	24 ABG71635 Human caspase recr
8	990	100.0	205	24 ABG74647 Human colon cancer
9	873.5	88.2	176	22 AA060594 Alternatively spliced

10	759	76.7	190	21 AA043675 Human cancer assoc
11	684	69.1	193	22 AA000592 Mouse target of me
12	684	69.1	193	22 AA020086 Mouse CARD-5 prote
13	684	69.1	193	23 AA093352 Mouse caspase recr
14	576	58.2	171	22 AA000593 Rat target of meth
15	564.5	57.0	158	22 AA074648 Human colon cancer
16	545	55.1	110	22 AA000595 Alternatively spliced
17	509	51.4	136	21 AA053881 Human colon cancer
18	448	44.3	91	22 AA000589 Human target of me
19	439	44.3	84	22 AA000591 Human target of me
20	305	30.8	76	23 AA021934 PAAD domain-conta
21	267	27.0	89	22 AA036608 Human FLEXIT-30 pr
22	267	27.0	89	23 AA017852 Pyrin domain conta
23	253	25.6	1397	22 AA072670 Human NB-ARC and C
24	250.5	25.3	1429	22 AA062571 Human CARD-7 poly
25	250.5	25.3	1429	23 ABG97969 Human leucine rich
26	250.5	25.3	1429	23 ABG78472 Leucine-rich repea
27	250.5	25.3	1429	23 ABG77916 Human leucine-rich
28	250.5	25.3	1429	24 ABG71631 Human caspase recr
29	250.5	25.3	1429	24 ABG71633 Human caspase recr
30	250	25.3	1442	22 AA072671 Human NB-ARC and C
31	250	25.3	1473	22 AA072669 Human NB-ARC and C
32	250	25.3	1473	22 AA072711 Human NAC beta iso
33	249.5	25.2	1429	23 ABG78455 Human caspase recr
34	249.5	25.2	1429	23 AA017855 Human caspase recr
35	248.5	25.1	1473	22 AA060758 Human G-protein co
36	240	24.2	442	21 AA024513 Human secreted pro
37	182	18.4	76	23 AA021935 PAAD domain-conta
38	153	15.5	65	21 AA024519 Human secreted pro
39	147	14.8	77	22 AA07516 Human PYRIN-1 doma
40	137	13.8	77	22 AA07515 Human PYRIN-1 doma
41	125.5	12.7	891	22 AA04546 Human G-protein co
42	125.5	12.7	1851	22 AA000023 Human activated T-
43	125.5	12.7	1851	23 AA017860 Pyrin domain conta
44	122.5	12.4	781	20 AA090001 FMF associated pro
45	122.5	12.4	781	23 AA017853 Pyrin domain conta

ALIGNMENTS

RESULT 1
ID AA198553 standard; Protein; 195 AA.
AC AA198553;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human breast tumour-associated protein 14.
XX
KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;
KM treatment; tumour; cytoskeletal; medicament.
XX
OS Homo sapiens.
XX
XX DB19813839-A1.
XX
XX 23-SEP-1999.
XX
XX 20-MAR-1998; 98DE-1013839.
XX
XX 20-MAR-1998; 98DE-1013839.
XX
XX (META-) METAGEN GDS GENOMFORSCHUNG MBH.
XX
XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX WPI; 1999-528981/45.
XX DR N-PSDB; AA233631.
XX
XX Human nucleic acid sequences and protein products from tumor breast
PT tissue, useful for breast cancer therapy -

XX Claim 22, 149; 188pp; German.
 CC This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AA48540-448617 represent protein
 CC fragments encoded by the expressed sequence tags described in the method
 CC of the invention.
 XX
 SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 20; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLREGYGRIPRGALLSMIDLDTDKLVSYF 60
 DB 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLREGYGRIPRGALLSMIDLDTDKLVSYF 60
 QY LETYGAELTANVLRDMLGLOEMAGOLQAATHQSGAAPAGIQAPPOGAARGLHFIIDQHRA 120
 DB 61 LETYGAELTANVLRDMLGLOEMAGOLQAATHQSGAAPAGIQAPPOGAARGLHFIIDQHRA 120
 QY 121 ALIARVTNVEMLLDALYGVLTDBOYQAVRAEPTNPSKKRKLFSFTPANMWTCKDILLQA 180
 DB 121 ALIARVTNVEMLLDALYGVLTDBOYQAVRAEPTNPSKKRKLFSFTPANMWTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195
 DT 16-JAN-2002 (first entry)
 XX Human novel cytokine encoded by cDNA 790CIP2B_1 #1.
 DE Human novel cytokine encoded by cDNA 790CIP2B_1 #1.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation;
 KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200175093-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US10484.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 22-SEP-2000; 2000US-0668680.
 PR 23-OCT-2000; 2000US-0695618.
 PR 30-NOV-2000; 2000US-0728711.

PR 14-MAR-2001; 2000US-0728711.
 XX
 XX (HYSE-) HYSEBQ INC.
 XX
 PI Tang YT, Aundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;
 PI Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-626432/72.
 DR N-PSDB; AAS59817.
 XX
 PT New polypeptides and nucleic acids, useful for diagnosis, treatment of
 PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
 PT degenerative disorders, cancer and promoting wound healing
 PT
 PS Claim 20; Page 242-243; 336pp; English.
 XX
 XX The invention relates to isolated human polypeptides (which may be
 CC cytokines) and the polynucleotides encoding them. The protein is useful
 CC for identifying a compound which binds to it (e.g. modulators, agonists
 CC and antagonists). The polynucleotides are useful as an array for mismatch
 CC detection. The proteins and nucleic acids are useful as nutritional
 CC sources or supplements. The protein exhibits activity relating
 CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity, immune stimulating or immune
 CC suppressing and activin or inhibin related activities. The proteins (and
 CC antibodies raised against them) and nucleic acids are therefore useful in
 CC the diagnosis and treatment of diseases and disorders such as cancer,
 CC central and peripheral nervous system diseases and neuropathies,
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
 CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
 CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth, and in tissue repair, healing of burns, incisions, ulcers, for
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
 CC periodontal disease, lung or liver fibrosis, reperfusion injury in
 CC various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
 CC such as asthma or other respiratory problems, coagulation disorders,
 CC haemophilia, septic shock, sepsis, arthritis, nephritis and inflammatory
 CC bowel disease, viral infection and are useful in altering bodily
 CC characteristics. The present sequence represents a novel protein of the
 CC invention.
 XX
 SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 22; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLREGYGRIPRGALLSMIDLDTDKLVSYF 60
 DB 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLREGYGRIPRGALLSMIDLDTDKLVSYF 60
 QY LETYGAELTANVLRDMLGLOEMAGOLQAATHQSGAAPAGIQAPPOGAARGLHFIIDQHRA 120
 DB 61 LETYGAELTANVLRDMLGLOEMAGOLQAATHQSGAAPAGIQAPPOGAARGLHFIIDQHRA 120
 QY 121 ALIARVTNVEMLLDALYGVLTDBOYQAVRAEPTNPSKKRKLFSFTPANMWTCKDILLQA 180
 DB 121 ALIARVTNVEMLLDALYGVLTDBOYQAVRAEPTNPSKKRKLFSFTPANMWTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195
 DT 16-JAN-2002 (first entry)
 XX Human novel cytokine encoded by cDNA 790CIP2B_1 #1.
 DE Human novel cytokine encoded by cDNA 790CIP2B_1 #1.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation;
 KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200175093-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US10484.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 22-SEP-2000; 2000US-0668680.
 PR 23-OCT-2000; 2000US-0695618.
 PR 30-NOV-2000; 2000US-0728711.

AC AAE00588;
 XX
 DT 02-JUN-2001 (first entry)
 XX
 DE Human target of methylation-induced silencing-1 (TMS1) protein.
 XX
 KW Human; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast.
 XX
 OS Homo sapiens.
 XX
 PN WO200129235-A2.
 XX
 PD 26-APR-2001.
 XX
 PF 18-OCT-2000; 2000WO-US28747.
 XX
 PR 18-OCT-1999; 99US-0159975.
 XX
 PA (UYEM-) UNIV EMORY.
 XX
 PI Vertino PM;
 XX
 DR WPI; 2001-290922/30.
 XX
 DR N-PSDB; AAD03889, AAD03890.
 XX
 PT Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis -
 XX
 XX
 XX Claim 85; Page 114; 124pp; English.
 XX
 CC The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
 CC abnormally low levels of TMS1 expression products. Unique fragments of
 CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by
 CC increasing TMS1 polypeptide level to an above normal level. The CpG
 CC island region of TMS1 or its fragments are used to study the methylation
 CC patterns apart from any coding region contained in it.
 CC The present sequence is human target of methylation-induced silencing-1
 CC (TMS1) protein.
 CC
 XX
 XX Sequence 195 AA;
 SQ
 Query Match 100.0%; Score 990; DB 22; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101; Indels 0; Gaps 0;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLRGGYGRIRGALLSMDALDLTDKLVSYFY 60
 DB 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLRGGYGRIRGALLSMDALDLTDKLVSYFY 60
 QY 61 LETYGAELITANVLDLMDGLQEMAGLOAATHOGSGAAPGIGIAPPQSAKPKLAHTIDGHR 120
 DB 61 LETYGAELITANVLDLMDGLQEMAGLOAATHOGSGAAPGIGIAPPQSAKPKLAHTIDGHR 120
 QY 121 ALIARVTNVEWLLDALYGVKVLTDGQYQAVRAEPTNPSTGRKLFSTTPAMNNTCKDILLQA 180
 DB 121 ALIARVTNVEWLLDALYGVKVLTDGQYQAVRAEPTNPSTGRKLFSTTPAMNNTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 4
 AAB20085
 ID AAB20085 standard; Protein; 195 AA.
 XX
 AC AAB20085;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Human CARD-5 protein.
 XX
 KW CARD-5; caspase recruitment domain; human; cancer; infection;
 KW autoimmune disease; neurological disease; haematological disease;
 KW immune disease; inflammation; antitumour; antiseptic;
 KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200100826-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000WO-US17691.
 XX
 PR 28-JUN-1999; 99US-0340620.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Bertin J;
 XX
 DR WPI; 2001-061973/07.
 XX
 DR N-PSDB; AAF30007.
 XX
 PT Isolated intracellular proteins predicted to be involved in regulating
 PT caspase activation are used for diagnosis and treatment of e.g. cancer,
 PT viral infections, autoimmune diseases, neurological diseases and
 PT haematological disorders -
 XX
 XX
 XX Claim 9; Fig 21; 208pp; English.
 XX
 CC The present sequence is that of human caspase recruitment domain 5
 CC (CARD-5), an intracellular protein predicted to be involved in
 CC regulating caspase activation. The sequence is predicted from an
 CC isolated cDNA clone (see AAF30007). Methods of diagnosing and
 CC treating patients suffering from a disorder associated with an
 CC abnormal level or rate of apoptotic cell death, abnormal activity
 CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
 CC necrosis factor receptor complex or abnormal activity of a caspase
 CC involve administering a compound that modulates the expression or
 CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
 CC therapy methods. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD-3,
 CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
 CC cell survival and cell growth. They can also be used to screen
 CC drugs or compounds that modulate their activity or expression and
 CC to treat disorders associated with insufficient or excessive
 CC production of CARD-3, -4, -5 or -6 protein, or production of an
 CC aberrant protein.
 CC
 XX
 XX Sequence 195 AA;
 SQ
 Query Match 100.0%; Score 990; DB 22; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101; Indels 0; Gaps 0;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLRGGYGRIRGALLSMDALDLTDKLVSYFY 60
 DB 1 MGRARDAIIDLALNTLAEELKKFKLLSVPLRGGYGRIRGALLSMDALDLTDKLVSYFY 60

QY 61 LETYGELTANVLRDNGLOEMAGOLQATHTQSGAAPAGIOAPPOGAAPGHLFIIDQHR 120
 CC
 Db 61 LETYGELTANVLRDNGLOEMAGOLQATHTQSGAAPAGIOAPPOGAAPGHLFIIDQHR 120
 CC
 QY 121 ALIARTNVEMLLDALYGKVLTDQYQAVRAEPTNSKKRKLFSPFPANWTCCKDLLQA 180
 CC
 Db 121 ALIARTNVEMLLDALYGKVLTDQYQAVRAEPTNSKKRKLFSPFPANWTCCKDLLQA 180
 CC
 QY 181 LRESQSYLVEDLERS 195
 CC
 Db 181 LRESQSYLVEDLERS 195
 CC
 RESULT 5
 AAU99353
 ID AAU99353 standard; Protein; 195 AA.
 XX
 AC AAU99353;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human caspase recruitment domain-5 (CARD-5) protein.
 XX
 KW Human; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 111..195
 FT Domain /label= CARD_domain
 XX
 PN WO200244354-A2.
 XX
 PD 06-JUN-2002.
 XX
 PF 29-NOV-2001; 2001WO-US44894.
 XX
 PR 01-DEC-2000; 2000US-0728721.
 XX 24-APR-2001; 2001US-0841879.
 XX
 PA (MILL-) MILLENITUM PHARM INC.
 XX
 PI Bertin J;
 XX
 DR WPI; 2002-557538/59.
 XX N-PSDB; ABK87966.
 XX
 PT Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis
 XX
 PS Claim 22; Fig 3; 100p; English.
 XX
 CC The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteinyl aspartate-specific proteinases) are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CAPs and different
 CC subtypes of CAPs may confer binding specificity. CARD-5 is an

CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B
 CC (NF-kappaB) transcription factor pathway and binds the CAPs of
 CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
 CC activity and NF-kappaB activation, regulate cell growth and cell death
 CC and be used in gene therapy. The CARD-5 polypeptides are useful for
 CC identifying compounds which bind to them and modulate their activity and
 CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
 CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
 CC can be used to treat immune disorders such as chronic inflammatory
 CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
 CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
 CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
 CC (including tuberculosis and lepromatous leprosy) and in screening and
 CC detection assays. Modulators of CARD-5 activity or expression are also
 CC useful for treating autoimmune disorders, such as systemic lupus
 CC erythematosus and arthritis, cell depletion, neurological disorders,
 CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
 CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
 CC aplastic anaemia, myocardial infarction and stroke. The sequence
 CC presented is the human caspase recruitment domain-5 (CARD-5) protein.
 XX
 SQ Sequence 195 AA;
 XX
 Query Match 100.0%; Score 990; DB 23; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRRARDALIDALENTAEEIKKKFKLKLISVPIREGGRIPRGALISMDALDITDKLVSY 60
 Db 1 MRRARDALIDALENTAEEIKKKFKLKLISVPIREGGRIPRGALISMDALDITDKLVSY 60
 QY 61 LETYGELTANVLRDNGLOEMAGOLQATHTQSGAAPAGIOAPPOGAAPGHLFIIDQHR 120
 Db 61 LETYGELTANVLRDNGLOEMAGOLQATHTQSGAAPAGIOAPPOGAAPGHLFIIDQHR 120
 QY 121 ALIARTNVEMLLDALYGKVLTDQYQAVRAEPTNSKKRKLFSPFPANWTCCKDLLQA 180
 Db 121 ALIARTNVEMLLDALYGKVLTDQYQAVRAEPTNSKKRKLFSPFPANWTCCKDLLQA 180
 QY 181 LRESQSYLVEDLERS 195
 Db 181 LRESQSYLVEDLERS 195
 RESULT 6
 AA017854
 ID AA017854 standard; Protein; 195 AA.
 XX
 AC AA017854;
 XX
 DT 20-AUG-2002 (first entry)
 XX
 DE Pryn domain containing protein Pycard.
 XX
 KW Pryn domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antineumatic; antiasthmatic;
 KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; glomerular nephritis;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX
 OS Unidentified.
 XX
 PN WO200240668-A2.
 XX
 PD 23-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-EP12545.
 XX
 PR 15-NOV-2000; 2000DE-1056687.
 XX 30-NOV-2000; 2000DE-1059595.
 PR

XX PA (APOT-) APOTEC RES & DEV LTD.
 XX PI Techopp J, Martinon F;
 XX WPI; 2002-427093/45.
 DR N-PSDB; AAL47126.
 XX PT New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies
 XX PS Claim 5; Fig 1; 116pp; German.
 XX CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention.
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 23; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDIILDALENTLAEELKKFKLLSVPLRGYGRIRPGALLSMDALDITDKLVSPY 60
 Db 1 MGRARDIILDALENTLAEELKKFKLLSVPLRGYGRIRPGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAEITANVLRDMGLQEMAGLOAATHQSGGAAPAGIOAPQOSAAKPGIHFIDQHR 120
 Db 61 LETYGAEITANVLRDMGLQEMAGLOAATHQSGGAAPAGIOAPQOSAAKPGIHFIDQHR 120
 QY 121 ALIARVNVEMLDALYGVKLTDEYOQAVRAEPTNPSKMKLFSFTPAWMTCKDLLLOA 180
 Db 121 ALIARVNVEMLDALYGVKLTDEYOQAVRAEPTNPSKMKLFSFTPAWMTCKDLLLOA 180
 QY 181 LRESQSYLVEDLERS 195
 Db 181 LRESQSYLVEDLERS 195
 RESULT 7
 ID ABG71635 standard; Protein, 195 AA.
 XX AC ABG71635;
 XX DT 10-JAN-2003 (first entry)
 XX XX Human caspase recruitment domain-5 (CARD-5).
 XX Human, caspase activity; caspase recruitment domain-5; CARD-5;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudocytotoxicity; ICEBERG; cell growth; cell death; inflammation; CARD-7;
 KW leukæmia; caspase activation; cancer; follicular lymphoma;
 KW leukæmia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytotoxic; virucide; immunosuppressive;
 KW dermatological; nephrotic; neuroprotective; cardiac;
 KW caspase recruitment domain-7.
 OS Homo sapiens.

XX FN US2002128198-A1.
 XX XX
 XX PD 12-SEP-2002.
 XX PF 27-NOV-2001; 2001US-0996617.
 XX PR 28-JUN-1999; 99US-0340630.
 XX PR 27-OCT-1999; 99US-0428252.
 XX PR 15-AUG-2001; 2001US-0931071.
 XX PA (BERT/) BERTIN J.
 XX PI Bertin J;
 XX WPI; 2003-028967/02.
 DR N-PSDB; ABSS6032.
 XX PT Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
 PT CARD-7 and CARD-5 in presence of test compound, measuring their
 PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
 PT altered
 XX PS Disclosure; Fig 7; 43pp; English.
 XX CC The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudocytotoxicity) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, myocardial infarction, stroke, and
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC represents human CARD-5. The sequence encoding CARD-5 is used to
 CC identify cDNA encoding CARD-7.
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 24; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDIILDALENTLAEELKKFKLLSVPLRGYGRIRPGALLSMDALDITDKLVSPY 60
 Db 1 MGRARDIILDALENTLAEELKKFKLLSVPLRGYGRIRPGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAEITANVLRDMGLQEMAGLOAATHQSGGAAPAGIOAPQOSAAKPGIHFIDQHR 120
 Db 61 LETYGAEITANVLRDMGLQEMAGLOAATHQSGGAAPAGIOAPQOSAAKPGIHFIDQHR 120
 QY 121 ALIARVNVEMLDALYGVKLTDEYOQAVRAEPTNPSKMKLFSFTPAWMTCKDLLLOA 180
 Db 121 ALIARVNVEMLDALYGVKLTDEYOQAVRAEPTNPSKMKLFSFTPAWMTCKDLLLOA 180

Db 121 ALIARVTNVEMLDALYGVLTDEQYQAVRAEPTNSKMRKLSFTPANMTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 Db 181 LRESQSYLVEDLERS 195

RESULT 8
 AAG74647
 ID AAG74647 standard; Protein; 205 AA.
 XX AAG74647;
 AC
 XX
 XX 03-SEP-2001 (first entry)
 DT
 DE Human colon cancer antigen protein SEQ ID NO:5411.
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma.
 XX Homo sapiens.
 OS
 XX MO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000MO-US26524.
 PF
 XX 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI
 XX MPI; 2001-235357/24.
 DR N-PSDB; AAH34052.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 XX Claim 11; Page 7035-7036; 9803PP; English.

AAH3943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing CC inactive proteins or to supplement the patients own production of P. CC Additionally, N may be used to produce the colon cancer-associated P, CC by inserting the nucleic acids into a host cell and culturing the cell CC to express the proteins. N and P can be used in the prevention, diagnosis CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 CC and AAB77789 represent sequences used in the exemplification of the CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were CC missing at time of publication, meaning no sequences are present for CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 205 AA;
 SQ

Query Match 100.0%; Score 990; DB 22; Length 205;
 Best Local Similarity 100.0%; Pred. No. 3.4e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRADAIIDALNTLREBLKKFKKLLSVLRGGRIRGALLSMDALDITDKLVSYF 60
 Db 11 MGRADAIIDALNTLREBLKKFKKLLSVLRGGRIRGALLSMDALDITDKLVSYF 70

QY 61 LETYGAELTANVLRDMGLQEMAGCOLAATHQSGGAPACIQAPPOSAARKGLHFIDQHR 120
 Db 71 LETYGAELTANVLRDMGLQEMAGCOLAATHQSGGAPACIQAPPOSAARKGLHFIDQHR 130

QY 121 ALIARVTNVEMLDALYGVLTDEQYQAVRAEPTNSKMRKLSFTPANMTCKDILLQA 180
 Db 131 ALIARVTNVEMLDALYGVLTDEQYQAVRAEPTNSKMRKLSFTPANMTCKDILLQA 190

QY 181 LRESQSYLVEDLERS 195
 Db 191 LRESQSYLVEDLERS 205

RESULT 9
 AAE00594
 ID AAE00594 standard; Protein; 176 AA.
 XX AAE00594;
 AC
 XX
 XX 02-JUL-2001 (first entry)
 DT
 DE Alternatively spliced form of human TMS1 protein (lacking exon2).
 XX
 XX Human; target of methylation-induced silencing-1; TMS1; cytostatic;
 KM antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KM caspase-recruiting domain; CARD; cancer; breast.
 XX
 XX Homo sapiens.
 OS
 XX MO200129235-A2.
 PN
 XX 26-APR-2001.
 PD
 XX 18-OCT-2000; 2000MO-US28747.
 PF
 XX 18-OCT-1999; 99US-0159975.
 PR
 XX (UYEM-) UNIV EMORY.
 PA
 XX Vertino PM;
 PI
 XX MPI; 2001-290922/30.
 DR N-PSDB; AAD03906.
 XX
 XX Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 XX apoptosis -
 XX
 XX Claim 85; Page 123; 124pp; English.

The invention relates to identification of target of methylation-induced CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and CC plays a role in induction of apoptosis. TMS1 gene and protein are useful CC as tools for diagnosing and treating a subject at risk of developing CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or CC abnormally low levels of TMS1 expression products. Unique fragments of CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by CC increasing TMS1 polypeptide level to an above normal level. The CpG CC island region of TMS1 or its fragments are used to study the methylation CC patterns apart from any coding region contained in it.
 CC The present sequence is alternatively spliced form of human target of CC methylation-induced silencing-1 (TMS1) protein lacking exon2.

XX Sequence 176 AA;
 SQ

Query Match 88.2%; Score 873.5; DB 22; Length 176;
 Best Local Similarity 90.3%; Pred. No. 2.3e-88;
 Matches 176; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

QY 1 MGRARDAILDLENLTAEELKKFKLKLSPVLRGGRIPRGALLSMDALDITDKLVSFY 60
 DB 1 MGRARDAILDLENLTAEELKKFKLKLSPVLRGGRIPRGALLSMDALDITDKLVSFY 60
 QY 61 LETYGAEILTANVLDMGLQEWAGOLQAATHGSGAAPAGIOAPPOSAKPGIAHFIIDQHRA 120
 DB 61 LETYGAEILTANVLDMGLQEWAGOLQAATHGSGAAPAGIOAPPOSAKPGIAHFIIDQHRA 101
 QY 121 ALIARVTNVEWLDAKYKVLTDQYQAVRAEPTNPSKMRLLFSTPAMNWTCDLLIOA 180
 DB 102 ALIARVTNVEWLDAKYKVLTDQYQAVRAEPTNPSKMRLLFSTPAMNWTCDLLIOA 161
 QY 181 LRESQSYLVEDLERS 195
 DB 162 LRESQSYLVEDLERS 176

RESULT 10
 AAB43675
 ID AAB43675 standard; Protein: 190 AA.
 AC AAB43675;
 DT 08-FEB-2001 (first entry)
 DE Human cancer associated protein sequence SEQ ID NO:1120.
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaesthetic; antirheumatic; antithrombotic; antiviral;
 KW antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; noctropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 KM
 XX Homo sapiens.
 OS
 XX WO200055350-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US05882.
 PF
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR
 XX N-PSDB; AAC77884.
 PT Novel, isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -
 PT
 XX
 XX Claim 11; Page 1731-1732; 23522p; English.
 PS
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antithrombotic;
 CC antiinflammatory; antihypertensive; antiallergic; antibacterial;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of

CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 CC
 XX Sequence 190 AA;
 SQ
 Query Match 76.7%; Score 759; DB 21; Length 190;
 Best Local Similarity 95.6%; Pred. No. 1.2e-75;
 Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDAILDLENLTAEELKKFKLKLSPVLRGGRIPRGALLSMDALDITDKLVSFY 60
 DB 11 MGRARDAILDLENLTAEELKKFKLKLSPVLRGGRIPRGALLSMDALDITDKLVSFY 70
 QY 61 LETYGAEILTANVLDMGLQEWAGOLQAATHGSGAAPAGIOAPPOSAKPGIAHFIIDQHRA 120
 DB 71 LETYGAEILTANVLDMGLQEWAGOLQAATHGSGAAPAGIOAPPOSAKPGIAHFIIDQHRA 130
 QY 121 ALIARVTNVEWLDAKYKVLTDQYQAVRAEPTNPSKMR 160
 DB 131 ALIARVTNVEWLDAKYKVLTDQYQAVR-PSPTQAR 168

RESULT 11
 AAE00592
 ID AAE00592 standard; Protein: 193 AA.
 AC AAE00592;
 DT 02-JUL-2001 (first entry)
 DE Mouse target of methylation-induced silencing-1 (TMS1) protein.
 XX
 KW Mouse; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast.
 KM
 XX Mus musculus.
 OS
 XX WO200129235-A2.
 PN
 XX 26-APR-2001.
 PD
 XX 18-OCT-2000; 2000WO-US28747.
 PF
 XX 18-OCT-1999; 99US-0159975.
 PR
 XX (UYEM-) UNIV EMORY.
 PA
 XX Vertino PM;
 PI
 XX WPI; 2001-290922/30.
 DR
 XX N-PSDB; AAD03904.
 PT Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis -
 PT
 XX
 XX Claim 85; Page 120; 124p; English.
 PS
 CC The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or

CC abnormally low levels of TMS1 expression products. Unique fragments of
 CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by
 CC increasing TMS1 polypeptide level to an above normal level. The Cpg
 CC island region of TMS1 or its fragments are used to study the methylation
 CC patterns apart from any coding region contained in it.
 CC The present sequence is mouse target of methylation-induced silencing-1
 CC (TMS1) protein.
 CC
 XX Sequence 193 AA;
 SQ
 Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
 QY 1 MGRARDALIDALENTTAEELKFKFKLKLISVPLREYGRIPRGALLSMALDITDKLSYF 60
 DB 1 MGRARDALIDALENTSDELKFKFKMLTVQLREYGRIPRGALLQMDALDITDKLSYF 60
 QY 61 LETYGABLTANVLRDMGLOEMAGLOAATHQSGAAPGIGAPPOSAKPGHFIIDQRA 120
 DB 61 LESYGLHETMTVLRDMGLOELAEQLQ-TTKESGAVAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNEMLDALYKVLTDQYQAVRAEPTNSKMKLFSFTPANWTKCKLLQA 180
 DB 119 ALIARTEVDGVLDAHSGSVLTGQYQAVRAETTSQDKMKLFSFVSNWTKCKSLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPLYWMDLEQS 193
 RESULT 12
 AAB20086 standard; Protein; 193 AA.
 ID AAB20086
 AC AAB20086;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Mouse CARD-5 protein.
 XX
 KW CARD-5; caspase recruitment domain; mouse; cancer; infection;
 KW autoimmune disease; neurological disease; haematological disease;
 KW immune disease; inflammation; antitumor; antiseptic;
 KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
 KW gene therapy.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Domain 110..179
 FT /note="CARD"
 XX
 PN MO200100826-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000MO-US17691.
 XX
 PR 28-JUN-1999; 99US-0340620.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Bertin J;
 XX
 XX WPI: 2001-061973/07.
 DR N-PSDB; AAF30008.
 XX
 XX Isolated intracellular proteins predicted to be involved in regulating
 PT caspase activation are used for diagnosis and treatment of e.g. cancer,
 PT viral infections, autoimmune diseases, neurological diseases and
 PT haematological disorders -

XX Claim 9; Fig 19; 208bp; English.
 PS
 XX The present sequence is that of mouse caspase recruitment domain 5
 CC (CARD-5), an intracellular protein predicted to be involved in
 CC regulating caspase activation. The sequence is predicted from an
 CC isolated cDNA clone (see AAF30008). It shows 71.8% amino acid
 CC identity to human CARD-5 (see AAB20085). Methods of diagnosing
 CC and treating patients suffering from a disorder associated with an
 CC abnormal level or rate of apoptotic cell death, abnormal activity
 CC of the Fas/Apo-1 receptor complex, abnormal activity of the tumour
 CC necrosis factor receptor complex or abnormal activity of a caspase
 CC involve administering a compound that modulates the expression or
 CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
 CC therapy methods. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD-3,
 CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
 CC cell survival and cell growth. They can also be used to screen
 CC drugs or compounds that modulate their activity or expression and
 CC to treat disorders associated with insufficient or excessive
 CC production of CARD-3, -4, -5 or -6 protein, or production of an
 CC aberrant protein.
 CC
 XX Sequence 193 AA;
 SQ
 Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
 QY 1 MGRARDALIDALENTTAEELKFKFKLKLISVPLREYGRIPRGALLSMALDITDKLSYF 60
 DB 1 MGRARDALIDALENTSDELKFKFKMLTVQLREYGRIPRGALLQMDALDITDKLSYF 60
 QY 61 LETYGABLTANVLRDMGLOEMAGLOAATHQSGAAPGIGAPPOSAKPGHFIIDQRA 120
 DB 61 LESYGLHETMTVLRDMGLOELAEQLQ-TTKESGAVAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNEMLDALYKVLTDQYQAVRAEPTNSKMKLFSFTPANWTKCKLLQA 180
 DB 119 ALIARTEVDGVLDAHSGSVLTGQYQAVRAETTSQDKMKLFSFVSNWTKCKSLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPLYWMDLEQS 193
 RESULT 13
 AAU99352 standard; Protein; 193 AA.
 ID AAU99352
 AC AAU99352;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Mouse caspase recruitment domain-5 (CARD-5) protein.
 XX
 KW Mouse; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers

FT Domain 110..193
 /label= CARD_domain
 XX
 XX W0200244354-A2.
 XX PD 06-JUN-2002.
 XX PF 29-NOV-2001; 2001WO-US44894.
 XX PR 01-DEC-2000; 2000US-0728721.
 XX PR 24-APR-2001; 2001US-0841879.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Berlin U.
 XX
 XX WPI; 2002-557538/59.
 XX DR N-PSDB; ABR87964.
 XX
 PT Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis
 PT
 PS Claim 22; Fig 1; 100pp; English.
 XX
 CC The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteine1 aspartate-specific proteases) are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CARDS and different
 CC subtypes of CARDS may confer binding specificity. CARD-5 is an
 CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B
 CC (NF-kappaB) transcription factor pathway and binds the CARDS of
 CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
 CC activity and NF-kappaB activation, regulate cell growth and cell death
 CC and be used in gene therapy. The CARD-5 polypeptides are useful for
 CC identifying compounds which bind to them and modulate their activity and
 CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
 CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
 CC can be used to treat immune disorders such as chronic inflammatory
 CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
 CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
 CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
 CC (including tuberculosis and lepromatous leprosy) and in screening and
 CC detection assays. Modulators of CARD-5 activity or expression are also
 CC useful for treating autoimmune disorders, such as systemic lupus
 CC erythematosus and arthritis, cell depletion, neurological disorders,
 CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
 CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
 CC aplastic anaemia, myocardial infarction and stroke. The sequence
 CC presented is the mouse caspase recruitment domain-5 (CARD-5) protein.
 XX
 SQ Sequence 193 AA;
 Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
 QY 1 MGRARDALIDALENTLBELEKFKKLSVLRGCGRIPIPGALLISMDALDITKLVSYF 60
 DB 1 MGRARDALIDALENTLBELEKFKKLSVLRGCGRIPIPGALLISMDALDITKLVSYF 60
 QY 61 LETYGAELTANVLRLDMGJOEWAGLOAATHOGSGAAPGACIAPQSAKPGAFRIDOHA 120
 DB 61 LETYGAELTANVLRLDMGJOELAEQLO-ITKESGSAVAAASVPQSTNRKRG-HVVDGRQ 118
 QY 121 ALIARVTNVEWLLDALYGVKVLITDGOYAVRAEPTNPSKRRKLSFTPAMNWTCKOLLQA 180
 DB 119 ALIARVTNVEWLLDALYGVKVLITDGOYAVRAEPTNPSKRRKLSFTPAMNWTCKOLLQA 178

QY 181 LRESQSYVEDLERS 195
 DB 179 LKEIHPLYVMDLRS 193
 RESULT 14
 ID AAB00593 standard; Protein; 171 AA.
 AC AAB00593;
 DT 02-JUL-2001 (first entry)
 DE Rat target of methylation-induced silencing-1 (TMS1) partial protein.
 XX
 XX Rat; target of methylation-induced silencing-1; TMS1; cytostatic;
 XX antiproliferative; apoptosis inducer; gene therapy; CpG island;
 XX caspase-recruiting domain; CARD; cancer; breast.
 XX
 OS Rattus norvegicus.
 FH Key Location/Qualifiers
 FT Misc-difference 36
 FT /note= "Encoded by AGV"
 XX
 XX W0200129235-A2.
 XX PD 26-APR-2001.
 XX PF 18-OCT-2000; 2000WO-US28747.
 XX PR 18-OCT-1999; 99US-0159975.
 XX (UYEM-) UNIV EMORY.
 XX
 XX Vertino PM;
 XX
 XX WPI; 2001-280922/30.
 XX DR N-PSDB; AAD03905.
 XX
 PT Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis
 PT
 PS Claim 85; Page 121; 124pp; English.
 XX
 CC The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
 CC abnormally low levels of TMS1 expression products. Unique fragments of
 CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by
 CC increasing TMS1 polypeptide level to an above normal level. The CpG
 CC island region of TMS1 or its fragments are used to study the methylation
 CC patterns apart from any coding region contained in it.
 CC The present sequence is rat target of methylation-induced silencing-1
 CC (TMS1) partial protein.
 XX
 SQ Sequence 171 AA;
 Query Match 58.2%; Score 576; DB 22; Length 171;
 Best Local Similarity 67.1%; Pred. No. 2e-55;
 Matches 116; Conservative 18; Mismatches 37; Indels 2; Gaps 2;
 QY 23 FKXKLISVPLRECGYGRIPRGALLISMDALDITDKLVSYFYETTYGAELTANVLRLDMGJOE 82
 DB 1 FKXKLISVPLRECGYGRIPRGALLISMDALDITDKLVSYFYETTYGAELTANVLRLDMGJOE 80

QY 83 GQQAATHGSGAAPAGIQAPPSAKPGIHFIDQRAALIARVTNWEMLDLYGKVL 142
DB 61 EQLQKME-SGAVATATVPAGCTARTE-HFVDQHQALIAVTEVDGLDLALYGNVLT 118
QY 143 DEQYQAVRAEPTNPSKKRKLFSFTPAWNTCKDLLLOALRESQSYVEDLERS 195
DB 119 EGYQAVRAETTNQNKRLFSFAPAWNLTCKLFLHALKQTPYLVTDLEQS 171

RESULT 15
AAG74648
ID AAG74648 standard; Protein; 158 AA.
XX AAG74648;
AC AAG74648;
XX
DT 03-SEP-2001 (first entry)
XX
XX
DE Human colon cancer antigen protein SEQ ID NO:5412.
XX
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WQ200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-0826524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI, 2001-235357/24.
DR N-PSDB; AAH34053.
XX
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
PS Claim 11; Page 7036-7037; 9803p; English.
XX
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated P,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAG77789 represent sequences used in the exemplification of the
XX present invention.
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 158 AA;

Query Match 57.0%; Score 564.5; DB 22; Length 158;
Best Local Similarity 75.8%; Pred. No. 3; 4e-54;
Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 ARDAIIDLALNTAEELKKFKLTLVPLREGYGRIPGALLSMDALDLTDKLVSYFLFT 63
|||||

DB 13 ARDAIIDLALNTAEELKKF-----KLVSFYLET 41
QY 64 YGAEITANVLRDMGLQEMAGLOQAATHGSGAAPAGIQAPPSAKPGIHFIDQRAALI 123
DB 42 YGAEITANVLRDMGLQEMAGLOQAATHGSGAAPAGIQAPPSAKPGIHFIDQRAALI 101
QY 124 ARVTNWEMLDLYGKVLTDQYQAVRAEPTNPSKKR 160
DB 102 ARVTNWEMLDLYGKVLTDQYQAVR--PSPPTQAR 136

Search completed: January 29, 2004, 13:45:54
Job time : 9.48584 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 4.80296 Seconds

(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990
Sequence: 1 MGRARDALDLLENLTAEEL.....LLQLALRESQSYVEDLERS 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	250	25.3	1192	2	T17255 hypothetical prote
2	87	8.8	323	2	T49723 hypothetical prote
3	86.5	8.7	372	2	S23326 gene M2.2 protein
4	86.5	8.7	439	2	S75545 hypothetical prote
5	85	8.6	237	2	B55210 heli protein - Ana
6	85	8.6	1488	2	AH0615 cell division prot
7	84	8.5	292	2	C64699 proteinase IV (EC
8	83.5	8.4	542	2	A84229 sensory histidine
9	83	8.4	455	2	UC6051 trypsin-like prote
10	83	8.4	545	2	A87136 hypothetical prote
11	83	8.4	877	2	JN0772 glucan endo-1,3-be
12	82.5	8.3	446	2	S16308 translation elonga
13	82.5	8.3	522	2	D82195 phage replication
14	82	8.3	292	2	F71819 probable proteinas
15	82	8.3	788	2	A72330 phenylalanine-tRNA
16	81	8.2	792	2	A80704 pyruvate, water di
17	80.5	8.1	718	2	B81058 guanosine-3',5'-bi
18	80.5	8.1	725	2	B81819 CDA peptide synthet
19	80.5	8.1	2117	2	T36180 feruloyl-CoA synth
20	80	8.1	596	2	F78729 chromosome segrega
21	80	8.1	1233	2	T40059 hypothetical prote
22	79	8.0	350	2	E87714 serine endoprotei
23	79	8.0	455	2	C91142 serine endoprotei
24	79	8.0	455	2	F85987 alanyl-tRNA synthet
25	79	8.0	860	2	G82310 cell division prot
26	79	8.0	1486	2	C64832 kinesin-like cell
27	79	8.0	1486	2	G90754 hypothetical prote
28	79	8.0	1486	2	E85618 cell division prot
29	79	8.0	1534	2	JH0228

30	78.5	7.9	322	2	P87549 type IV secretion
31	78.5	7.9	346	2	A48470 translation elonga
32	78.5	7.9	424	2	AC2975 oxidoreductase ord
33	78.5	7.9	424	2	H98307 probable oxidoredu
34	78.5	7.9	699	2	T18984 hypothetical prote
35	78.5	7.9	956	2	B71468 probable insulinas
36	78.5	7.9	966	2	T51933 kinesin motor prot
37	78.5	7.9	4684	2	A59404 plectin [imported]
38	78	7.9	325	2	A57482 hypothetical prote
39	78	7.9	325	2	A12699 conserved hypothet
40	78	7.9	335	2	C75619 histidinol-phosphat
41	78	7.9	351	2	AF3419 phage host specific
42	78	7.9	365	2	G84849 probable actin [im
43	77.5	7.8	228	2	C87467 ABC transporter, A
44	77.5	7.8	542	2	U00661 heat shock protein
45	77.5	7.8	542	2	G32106 groEL protein - La

ALIGNMENTS

RESULT 1

T17255 hypothetical protein DKFP58601822.1 - human (fragment)

C.Species: Homo sapiens (man)

C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C.Accession: T17255

R.Koehrer, K.; Beyer, A.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A.Reference number: Z18722

A.Accession: T17255

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-1192 <KOE>

A.Cross-references: EMBL:AL117470

A.Experimental source: adult uterus; clone DKFP58601822

C.Genetics:

A.Note: DKFP58601822.1

Query Match 25.3%; Score 250; DB 2; Length 1192;
Best Local Similarity 44.5%; Pred. No. 1.7e-13;
Matches 65; Conservative 23; Mismatches 48; Indels 10; Gaps 4;

QY	54	DKLV-FILFRTYGMELTNV--LRDMGLQEMA-----GQLQATHTGSGAAPGICAPQS	106
DB	1036	DQLRSEFYVGHIGSGIRLQVXKDKDETLVWELVVPGLMPRT--TLIPPARIAVPSPL	1092
QY	107	AAKPLAFIDQRAALIRVTNVEWLDALYGVLTDEQYQAVRAEPTNPSKRLTFSFT	166
DB	1093	DAPQLHFTVDQYREGLIRVTSVEVYLDKLGQVLSQGYERVLAEINTRPSQMRGLFSL	1152
QY	167	PAMWTCDDLLQLALRESQSYVEDL	192
DB	1153	QSWDRCKDGLYQALMKETHPHLMEL	1178

RESULT 2

T49723 hypothetical protein B23L21.360 [imported] - Neurospora crassa

C.Species: Neurospora crassa

C.Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C.Accession: T49723

R.Schulte, U.; Aign, V.; Hohnsbeil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A.Reference number: Z25022

A.Accession: T49723

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-323 <SCH>

A.Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.360

A.Experimental source: BAC clone B23L21; strain OR74A

C.Genetics:

A.Gene: NCSP:B23L21.360

A:Map position: 6
A:introns: 183/1; 229/2

Query Match 8.8%; Score 87; DB 2; Length 323;

Best Local Similarity 20.7%; Pred. No. 5.1;
Matches 44; Conservative 41; Mismatches 80; Indels 48; Gaps 9;

QY 6 DALIDLENTAELEKFKKFKLKVLPREGYGRIPGALLSMALDLTDKLVSEFYLETYG 65
DB 119 DNFHDLKVCQGGDFEFQVSLP-----SSPTTIRFAFTENMAKHYRSRG 168
QY 66 AELTANVRDMGQEMAGQLOAATHGSGAPAGIOAPPOSAKPGIHFIDQRAALIA 125
DB 169 -----EMSVR-NEGIVVA--HAGSGTILEVLRQVLPVVPPTLMDNQAEL--- 213
QY 126 VINNVEMLDALYKGV--LTDEQYQAVRAEPTNPSKRRKLFSTP-----AMNV 171
DB 214 ADELSESTGAVVYRLGLTEAIRASLELVAGQGLKDKLPPSPPPVPSDSRYTLFDW 273
QY 172 ---TC-----KDLALRESQSYLVEDLERS 195
DB 274 MTLTCYPDLELRKQHLQDLNVEKSPQRELEQA 306

RESULT 3

gene ML2.2 protein precursor - Streptococcus pyogenes

C:Species: Streptococcus pyogenes
C:Date: 22-Nov-1993 #sequence_rev: 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S23326

R:Bessen, D.E.; Pischetti, V.A.
Infect. Immun. 60, 124-135, 1992

A:Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strep
A:Reference number: S23325; MUID:92104652; PMID:1370269

A:Accession: S23326

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-372 <BES>

A:Cross-references: EMBL:X61276; NID:947369; PIDN:CAA3582.1; PID:947371

C:Superfamily: M5 protein

Query Match 8.7%; Score 86.5; DB 2; Length 372;
Best Local Similarity 26.8%; Pred. No. 6.7;

Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

QY 12 LENLTAELKFKFKLKVLPREGYGRIPGALLSMALDLTDKLVSEFYLETYG 66
DB 78 LEKINAEERKFKLEKINENYKLDG---IDALEKEKEDTKTLAKTKENEIS 133
QY 67 ELTANVL-RDMGQEMAGQLOAATHGSGAPAGIOAPPO--SAKPGI-HFIDQRAAL 122
DB 134 EASRKGSLRDLASRTAKKLEAKHQKLEAKNKLTGQNVSAKSGLSNDLEASRAA- 192
QY 123 IAEVTNEMLDALYKGLTDEQ-----YQAVRAEPTNPSKRRKLFSTPAMNWTCKDL 177
DB 193 -----KKELEAKYKQLETDHQALEAKHQKLEADYGVSETSRKGLS-----RD-- 224
QY 178 LQALRESQSYLVEDLERS 195
DB 235 LEASREANKKVTSELTOA 252

RESULT 4

hypothetical protein slr0806 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.
C:Accession: S75545

C:Date: 25-Apr-1997 #sequence_rev: 25-Apr-1997 #text_change 20-Jun-2000

A:Variety: PCC 6803
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.
A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75545

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-439 <RAN>

A:Cross-references: EMBL:D90911; GB:AB001339; NID:91653083; PIDN:BA18106.1; PID:9165315

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: glycolate oxidase chain glcd

Query Match 8.7%; Score 86.5; DB 2; Length 439;
Best Local Similarity 24.0%; Pred. No. 8.2;
Matches 42; Conservative 22; Mismatches 60; Indels 51; Gaps 8;

QY 31 PAREGYGR---IPRGALLSMALDLTDKLVSEFYLETYGELTANVRDMGQEMAGQLOA 87
DB 194 PVHGYGTNGITTEITLPLTPALPMEAAVSR-----TNLSGALFAQN 237
QY 88 ATHQ--GSGAPAGIOAP-----QSAKPGIHFIDQRAALIAVTNVEM---TL 133
DB 238 LAHQDQIVKEISIQADPIPOYFSSLSKSYQGAHV-----MYVSELDWLAFTOLA 290
QY 134 DALYKGLTDEQYQAVRAEPTNPSKRRKLFSTPAMNWTCKDLLQALRESQSYL 188
DB 291 KASKGEIIEQ-----DPOSGKKINLIEF---NMHTTLIARAVDPSTLYL 334

RESULT 5

E55210
heci protein - *Anabaena* sp. (strain PCC 7120)

C:Species: *Anabaena* sp.
C:Date: 05-May-1995 #sequence_rev: 05-May-1995 #text_change 15-Oct-1999
C:Accession: E55210

R:Black, T.A.; Wolk, C.P.
J. Bacteriol. 176, 2282-2292, 1994

A:Title: Analysis of a Hec - mutation in *Anabaena* sp. strain PCC 7120 implicates a second
A:Reference number: A55210; MUID:94209228; PMID:8157596

A:Accession: E55210

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-237 <BLA>

A:Cross-references: GB:L22883; NID:9441119; PIDN:AAA22003.1; PID:9441121

A:Gene: heci

A:Start codon: TTG

Query Match 8.6%; Score 85; DB 2; Length 237;
Best Local Similarity 23.7%; Pred. No. 5.2;

Matches 49; Conservative 32; Mismatches 72; Indels 54; Gaps 12;

QY 14 NTAELKFKFKLKVLPREGYGRIPGALLSMALDLTDKLV---SYLETYGAEI 70
DB 12 NLT---LSDDEVHLMRIPLDQ-----PESQLDPLAATLSSDELARANRFYFPEHRRRFTA 63
QY 71 --NVLRDMGQEMAGQLOAATHGSGAPAGIOAPPOSAKPP-----GHAFIDQHR 119
DB 64 GRGILRSI---LGSYL-----GVEPGYKPYDSGKPIFGPRAESGLLFTLSHS 111
QY 120 AALIAAVTN-----VEML-----LDALYKGLTDEQYQAVRAEPTNPSKRRKLFST 166
DB 112 QNLACAVNVTQIGIDLEFLRPTSDLSIAKRFPLREYELRSIP--DEQKQKIFRY- 169
QY 167 PAMNWTCKDLLQALRESQSYLVEDLE 193
DB 170 ---WTCKEAYIKATGDGIAKL-EETL 191

RESULT 6

AH0615

cell division protein [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium

C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_rev: 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AH0615
 R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Conerton, F.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mout, S.; O'Geare, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AH0615
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1488 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD05394.1; PID:G16502156; GSPDB:GN00176
 C:Genetics:
 A:Gene: mukB
 C:Superfamily: Escherichia coli mukB protein

Query Match 8.6%; Score 85; DB 2; Length 1488;
 Best Local Similarity 24.0%; Pred. No. 50;
 Matches 42; Conservative 33; Mismatches 56; Indels 44; Gaps 11;

QY 3 RARDALD-----ALENLTAEELKKFKLKLSVP-----LRGGRIRPGALLS- 46
 DB 638 RREBAIVERDEVGARKVADEIER-----LSQCGAEDQRLNALAEFRG---GVLLSE 688

QY 47 -MDALDITDLKVSRYLETYGAEELTANVLRDMGLQEMAGCOLAATHQSGAAPAG---TQA 102
 DB 689 IYDDVSLD--APFSAIYGPSRAIIVPD--LSQIAEOLGGLTD-----CPEDLYIEG 739

QY 103 PPOSAKPGHFIIDQRAALIAIARTVNEWLLD-----ALYKVLTDROYQAVRAE 152
 DB 740 PDQS-FDQSVFSDLEKAVVVKIADRKMRSPSPITGRARAKRISLHAE 793

RESULT 7
 C64699
 protease IV (EC 3.4.-.-) - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C>Date: 09-Aug-1997 #sequence_revision 15-Aug-1997 #text_change 29-Sep-1999
 C/Accession: C64699
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, P.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Meldman, J.M.; Fujii, C.; Bowman, C.; Wathey, L. Nature 388, 539-547, 1997
 A:Authors: Wallin, B.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*. A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: C64699
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-292 <TOM>
 A:Cross-references: GB:AE000644; GB:AE000511; NID:92314609; PIDN:AAD08477.1; PID:G231461
 C:Superfamily: proteinase IV MTH806
 C:Keywords: hydrolase

Query Match 8.5%; Score 84; DB 2; Length 292;
 Best Local Similarity 23.6%; Pred. No. 8.2;
 Matches 41; Conservative 30; Mismatches 61; Indels 42; Gaps 7;

QY 42 GALLSMALDL-TDKL-----VSFYLETYGAEELTANV-----LRDMGLQEMAGLOA 87
 DB 57 GALTSTDFDKXVDKILKTSIKVLLILSPGAVASVLSKIDL-KQKVPVLAIA 115

QY 88 ATHQSGAAPAGIQAPQSAKPGIHFIDQRAALIAIARV-----TNVWLLDALYKVL 141
 DB 116 RGVASGSYYAGMQASRYVASK-----ASLISIGIVFSGAVENVILNKV----- 160

QY 142 TDEYQAVRAEPTNPSTKRLFTTPAMNTCKULLQALRESQSYVEDLERS 195
 DB 161 -----GVATQGVAGGEYKEIGTPTRAKPKNEKDFLNTLVNEYQYQVFNVAKA 208

RESULT 8

A84229
 sensory histidine protein kinase homolog [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: A84229
 R:Ne, M.V.; Kennedy, S.P.; Maharis, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor Jung, K.H.; Alam, M.; Preltse, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniele, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A:Title: Genome sequence of Halobacterium species NRC-1. A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: A84229
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-542 <STO>
 A:Cross-references: GB:AE004437; NID:G10580295; PIDN:AAG19197.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: atfA2

Query Match 8.4%; Score 83.5; DB 2; Length 542;
 Best Local Similarity 28.2%; Pred. No. 19;
 Matches 37; Conservative 18; Mismatches 53; Indels 23; Gaps 6;

QY 37 GRIPRALLSMDAL--DLTDKVSFYLE-----TYGAELTANVLRDM--GLQEMAGLOA 88
 DB 40 GRIEVANEARADVIGADARLVETFFWQAPFTHDDQVAAREQVAAVQGEASQF-TA 98

QY 89 THQSGAAPAGIQAPQSAKPGIHFIDQ-----RAALIAIARTVNEWLLD--- 134
 DB 99 THHTSGGRTAVLELQPMAPAPAVDGIIDHDHAFVAVVGRRAASSTADVESAIDAVQ 158

QY 135 ALYKVLTDROYQ 145
 DB 159 ALYATATDQ 169

RESULT 9
 JC6051
 trypsin-like proteinase (EC 3.4.21.-) hhoA precursor - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 01-Mar-2002
 C/Accession: JC6051
 R:Bass, S.; Gu, Q.; Christen, A.
 J. Bacteriol. 178, 1154-1161, 1996
 A:Title: Multicopy suppressors of Prc mutant *Escherichia coli* include two HtrA (DegP) prc A:Reference number: JC6051; MUID:96165273; PMID:8576052
 A:Accession: JC6051
 A:Molecule type: DNA
 A:Residues: 1-455 <BAS>
 A:Cross-references: GB:U15661; NID:9558911; PIDN:AAC43992.1; PID:G558912
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co A.; Rose, D.J.; Mau, B.; Shaoh, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12. A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-455 <BLAT>
 A:Cross-references: GB:AE000402; GB:U00096; NID:G1789619; PIDN:AAC76266.1; PID:G1789629; A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: hhoA
 C:Superfamily: Helicobacter serine proteinase
 C:Keywords: hydrolase; serine proteinase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-455/Product: trypsin-like serine protease hhoA #status predicted <MAT>
 F:109,139,214/Active site: His, Asp, Ser #status predicted

Query Match 8.4%; Score 83; DB 2; Length 455;
 Best Local Similarity 21.9%; Pred. No. 17;
 Matches 55; Conservative 34; Mismatches 76; Indels 86; Gaps 10;

```

QY 11 ALENTAEELKKFKKLLSVPLREYGRIPRG-ALISMDALDITDKVSEF-----YLET 63
DB 217 ALININNGE-----LIGINPAIILAPGGGSVIGFAIPSNARITLAQQLIDFGEIKKGLGI 271
QY 64 YGAELTANVLRLDMGLQEMAGLOAATHQSGAAPADIOA----- 102
DB 272 KTEHSADIAKAFNVLQKAFVSEVLPGSSAKAGVAKGDIITTSUNGKPLNSPFLRSR 331
QY 103 ----PPOSAAKQGLHFIDQRAALIRAVT-----NVEMLLDALYGVLTDEQY--- 146
DB 332 IATTEGTGVKIGLL---RNGKPLEVEVLTIDISTSSSAEMITPLLEGATISDGLKKG 388
QY 147 -----QAVRAEP-----TNPSSKPKLFSTPAMNWTCKDLILQ 179
DB 389 GKGIKIDDEVVKSPPAQAQLOKQDVIIIGVNRDRVNSIAEMRKVLAAPKPA-----IALQ 442
QY 180 ALRESQS-VLV 189
DB 443 IYRGNESTIYLL 453

```

RESULT 10

hypothetical protein ML1816 [imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C/Accession: A87136
R/Col: S.T.; Eizlmeier, K.; Parhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho R.; Davies, R.M.; Devlin, K.; Duhoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq A/Title: Massive gene decay in the leprosy bacillus.
A/Reference number: A86909; PMID:21128732; PMID:11234002
A/Accession: A87136
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-545 <STO>
A/Cross-References: GB:AL450380; NID:G13093526; PIDN:CAC30769.1; GSPDB:GN00147
C/Genetics:
A:Gene: ML1816
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 8.4%; Score 83; DB 2; Length 545;
Best Local Similarity 24.4%; Pred. No. 22;
Matches 41; Conservative 15; Mismatches 50; Indels 62; Gaps 7;
QY 51 DLTDKIVSYF-----LETYGALTA-NVLRDMGLQ-----EM 81
DB 113 DARDRAIRRYGQLEERFLALGYSASESGRICASGLPERVLVQQLCTLSGGQRRRVEL 172
QY 82 AGQLQATHQSGAAPAG-----IQAPQSAKPGILHF-----IDQRAALIA 124
DB 173 AILPLPASAAGTCASSGCTTLLDEPTNHLDSLSGLMDPLRSHGTGLVISHNVELIA 232
QY 125 RVTNVEMLDALYGVK-----LTDEYQAVRAEPTNSK 158
DB 233 AVNVRWFIDAVLVKGVYVMGWYKYLDSRADTEQRR--RRERVNER 278

RESULT 11

glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) precursor - Bacillus circulans (strain N/Alternate names: beta-1,3-glucanase

C/Species: Bacillus circulans
C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C/Accession: JN0772; P0613
R/Col: M.; Aono, R.; Horikoshi, K.
Bioesci. Biotechnol. Biochem. 57, 1518-1525, 1993
A/Title: Structure of the 87-kDa beta-1,3-glucanase gene of Bacillus circulans. JAM165
A/Reference number: JN0772; PMID:94033858; PMID:7764221
A/Accession: JN0772

A/Molecule type: DNA
A/Residues: 1-877 <YAM>
A/Cross-References: GB:D17519; NID:9601877; PIDN:BA04469.1; PID:9601878
A/Accession: P0613
A/Molecule type: protein
A/Residues: 39-58 <YAA>

A/Genes: bglH
C/Keywords: glycosidase; hydrolase; polysaccharide degradation
F/1-38/Domains: signal sequence #status predicted <SIG>
F/39-877/Product: glucan endo-1,3-beta-D-glucosidase #status predicted <MAT>

Query Match 8.4%; Score 83; DB 2; Length 877;
Best Local Similarity 22.3%; Pred. No. 39;
Matches 40; Conservative 26; Mismatches 73; Indels 40; Gaps 8;

QY 10 DALENTAEELKKFKKLLSVPLR-----EGYGRIPRGALISMDALDITDKVSEFLE-- 62
DB 471 DRAQNVFVQD-GKNIKALNEPKSPPOPSRYAQISSGKINKDKHFLSKYGRVDRAPKLP 529
QY 63 -----TYGAELTANVLRLDMGLQEMAGLOAATHQSGAAPAGIQAPQS 106
DB 530 TNGGIWPLMLPQDNVYGTWASSG---EIDVWEAKGRLPGST---SGAVHFGQWPTNR 583
QY 107 AAKPGILHF-----IDQRAALIRAVTNVEMLLDA-LYGVLTDEYQAVRAEPTNS 157
DB 584 YLSGEYHFPEEGTFANDYHVSVWEEDNIKWYDGKFFKVTROQWYSA--AAPNNPN 640

RESULT 12

translational elongation factor eEF-1 alpha chain - Stylyomychia lemnae

C/Species: Stylyomychia lemnae
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
C/Accession: S16308
R/Bierbaum, P.; Doenhoff, T.; Klein, A.
Mol. Microbiol. 5, 1567-1575, 1991
A/Title: Macromolecular and microstructural configurations of a gene encoding the protein syn A/Reference number: S16308; PMID:9215784; PMID:1840642
A/Accession: S16308
A/Molecule type: DNA
A/Residues: 1-446 <BIB>
A/Cross-References: EMBL:X57926; NID:G10139; PIDN:CAA41001.1; PID:G10140
A/Note: the authors translated the codon GAT for residue 307 as Gly
C/Genetics:
A:Gene: efaA
A/Genetic code: SGC5
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
C/Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis
F/8-156/Domains: translation elongation factor Tu homology <ETU>
F/14-21/Region: nucleotide-binding motif A (P-loop)
F/153-156/Region: GTP-binding NKXD motif

Query Match 8.3%; Score 82.5; DB 2; Length 446;
Best Local Similarity 21.4%; Pred. No. 19;
Matches 44; Conservative 40; Mismatches 81; Indels 41; Gaps 8;

QY 8 IUDALENTAEELKKFKKLLSVPLREY-----GRIPRGAL-----LSMDALD 52
DB 216 LIDALDAL--DQKPKPKPLPLPLQDYKIGIGTVPGRVETGLLRGWMLTPAPNNI 273
QY 53 TDKIVSYFLETYGALTA-----NV---LRDMGLQEMAGLOAATHQSGA-----AP 97
DB 274 TTECKS--VEMHESLTEAEPGDNGFTVKNLSVLDLRGVYASQKNDPAKOTTNFLQ 331
QY 98 AGIQAPQSAKPGILHFIDQRAALIRAVTNVEMLLDALYGVLTDE-----QYAVR 150
DB 332 VAVLHNPQIQGVAPVLDCHTAHACKFDEHSKVDKRSKVLDEEPFIRISGGAALVR 391
QY 151 AEPNPSKPKKLFSTPAMNWTCKDL 176
DB 392 WPKQKMCVEAENQYPPILGRFAVRDM 417

RESULT 13

DB2195

Phase replication protein Cri VC1469 VC1475 [imported] - *Vibrio cholerae* (strain N16961)C/Species: *Vibrio cholerae*

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: DB2195; A82196

R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A/Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: DB2195

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE2>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94631.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

A/Accession: A82196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <HE1>

A/Cross-references: GB:AE004225; GB:AE003852; NID:96555968; PIDN:AAF94626.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

Matches 40; Conservative 32; Mismatches 60; Indels 42; Gaps 7;

QY 42 GALLSDADL-TDKL-----VSFYLETYAEITANV-----LRDMGLQEMAGLOA 87

D 57 GAFSTEDFEKVDKLTPTSGVILLIDSPGASASVSEKXADL-KQKPPVLA 115

QY 88 ATHQSGAAPAGIQAPPOSAAKPGHFIHQRAALIAV-----TNVEMLDALYKVL 141

D 116 KGVMAAGSYIAQQAEBVYASK-----ASLIGSIVITSSANVENLTKV----- 160

QY 142 TDEQYAVRAEPTNPKRMKLFSPFPAWMTCKDLLLQALRESQSYLVEDLES 195

D 161 -----GVATQGVHAGYKXKIGFTPAKMKNEKFLQNLVNEQYQWVDDVAKA 208

RESULT 15

A72330

phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain - *Thermotoga maritima* (strain MSB8)C/Species: *Thermotoga maritima*

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002

C/Accession: A72330

R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A/Reference number: A72200; MUID:9287316; PMID:10360571

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-788 <ARN>

A/Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135;

A/Experimental source: strain MSB8

A/Accession: A72330

QY 66 AELTANVLRDMGLQEMAGLOAATHQ-----GSGAAPAGIQAPPOSAAKPGHFIHQ 117

D 140 VELGVNVVEEYGLNERVLDIEITPNRPDCLSIIIGVARELSALSGRPLNKQPDVVSFYDE 198

Search completed: January 29, 2004, 13:49:47
Job time : 6.80296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 / Search time 3.00185 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARPAIDALENTAEEL.....LLQALRESQSYVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	1 ASC_HUMAN	Q9ULZ3 homo sapien
2	684	69.1	193	1 ASC_MOUSE	Q9EPB4 mus musculu
3	250	25.3	1473	1 NALI_HUMAN	Q9C000 homo sapien
4	213	21.5	203	1 ASC_ERAB	Q91966 brachydanio
5	138	13.9	854	1 PYAS_RAT	Q63035 rattus norv
6	132	13.3	843	1 PYAS_MOUSE	Q61W62 mus musculu
7	125.5	12.7	892	1 MEVF_HUMAN	P59044 homo sapien
8	122.5	12.4	781	1 MEVF_HUMAN	O15553 homo sapien
9	99.5	10.1	980	1 PYA3_HUMAN	Q9W494 homo sapien
10	99.9	10.0	431	1 CAR8_HUMAN	Q9Y262 homo sapien
11	91.5	9.2	143	1 SILE_SALTY	Q9Z4T3 salmoneila
12	89	9.0	338	1 RUVB_THETN	Q9Z4T3 thermococc
13	86.5	8.7	372	1 M22_STRPY	P50469 streptococc
14	86.5	8.7	835	1 SYFB_CORGL	Q9N966 corynebacte
15	86	8.7	1033	1 CISI_MOUSE	Q9Z4B8 mus musculu
16	85	8.6	237	1 HETI_AANSP	P37695 anabesna sp
17	84.5	8.5	340	1 RIK2_HUMAN	O43353 h recepto
18	83	8.4	378	1 ACT_SCHDU	O65314 scheriffelia
19	83	8.3	455	1 DEQO_ECOLI	P39099 escherichia
20	82.5	8.3	446	1 EPJA_STYLE	P25166 stylyonchia
21	82	8.3	788	1 SYFB_THEMEA	Q9W494 thermocoga
22	80	8.1	539	1 RIK2_MOUSE	P37695 mus musculu
23	80	8.1	1034	1 CISI_HUMAN	Q9Z4B8 mus musculu
24	80	8.1	1233	1 SMCI_SCHPO	Q94383 schizosacch
25	79	8.0	860	1 SYA_VIBCH	O66648 vibrio chol
26	79	8.0	1486	1 MUKB_ECOLI	P22523 escherichia
27	78.5	7.9	346	1 EPJA_EIMBO	Q97051 elmeria bov
28	78.5	7.9	4684	1 PLEI_HUMAN	O15148 homo sapien
29	78	7.9	366	1 ACT9_ARATH	P33733 arabidopsis
30	76.5	7.7	324	1 PARB_TREPA	Q9X016 trepomonas p
31	76.5	7.7	953	1 CAR4_HUMAN	Q9Y223 homo sapien
32	76.5	7.7	739	1 PURL_LISMO	Q9Y661 listeria mo
33	76	7.7			

ALIGNMENTS

RESULT 1	ASC_HUMAN	STANDARD	PRT	195 AA.
ID	ASC_HUMAN	Q9ULZ3; Q9BD12; Q9BS25; Q9HBD0; Q9NXJ8;		
AC	Q9ULZ3; Q9BD12; Q9BS25; Q9HBD0; Q9NXJ8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Apoptosis-associated speck-like protein containing a CARD (hASC)			
DE	(PYCARD) (Target of methylation-induced silencing 1) (Caspase recruitment domain protein 5).			
GN	ASC OR TW51 OR CARD5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Leukemia;			
RX	MEDLINE=20036508; PubMed=10567338;			
RA	Maunoto J., Taniguchi S., Ayukawa K., Sarvorham H., Kishino T., Nishikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J., "ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells.";			
RT	J. Biol. Chem. 274:33835-33838(1999).			
RL	[2]			
RN	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RP	TISSUE=Fibroblast;			
RC	MEDLINE=20552139; PubMed=11103776;			
RX	Conway K.E., McConnell B.B., Bowring C.E., Donald C.D., Warren S.T., Vertino P.M.;			
RA	"TW51, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers.";			
RT	Cancer Res. 60:6236-6242(2000).			
RL	[3]			
RN	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	Marillon F., Hofmann K., Tschopp J.;			
RA	"PyCard, a PYD and CARD containing molecule.";			
RT	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
RL	[4]			
RN	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	Bertin J.;			
RA	"CARDs protein is a CARD/PYRIN family member that is involved in apoptosis signal transduction.";			
RT	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RL	[5]			
RN	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	TISSUE=Colon mucosa;			
RC	Watanabe K., Kunitake A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;			
RA	"NEDD human cDNA sequencing project.";			
RT	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RL	[6]			
RN	SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).			
RP	TISSUE=Lymph, and Pancreas;			

34	75.5	7.6	444	1	TIG_AZOR	O9X6W7 azospirillum
35	75.5	7.6	542	1	CH60_LACIA	P37282 lactococcus
36	75.5	7.6	542	1	CH60_LACLC	Q9AEP7 lactococcus
37	75.5	7.6	546	1	CH60_LACAC	Q93G07 lactobacilli
38	75.5	7.6	611	1	BIR_CHICK	O90660 gallus galli
39	75.5	7.6	1698	1	Y076_HUMAN	Q14939 homo sapien
40	75	7.6	545	1	CH60_PARDE	O9Z462 paracoccus
41	75	7.6	791	1	PPSA_ECOLI	P23538 escherichia
42	74.5	7.5	401	1	ISPE_LICES	P93841 lycoperisico
43	74.5	7.5	862	1	LOXA_PPAUV	P27480 phaseolus v
44	74.5	7.5	957	1	SECA_MYCSM	P71533 mycobacteri
45	74.5	7.5	1062	1	PYA7_HUMAN	P59046 homo sapien

RA MEDLINE=22388257; PubMed=12477932;
 RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist P.,
 RA Diatchenko L., Marusik K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schlez T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley D.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=20552140; PubMed=11103777;
 RA McConnell B.B., Vertino P.M.;
 RT "Activation of a caspase-9-mediated apoptotic pathway by subcellular
 RT redistribution of the novel caspase recruitment domain protein TWS1.";
 RL Cancer Res. 60:6243-6247(2000).
 CC -1- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
 CC activity is mediated predominantly through the activation of
 CC caspase 9.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
 CC a redistribution from the cytoplasm to the aggregates occurs.
 CC These appeared as hollow, perinuclear spherical, ball-like
 CC structures.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9ULZ3-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9ULZ3-2; Sequence=VSP_004119;
 CC Name=3;
 CC IsoId=Q9ULZ3-3; Sequence=VSP_004118;
 CC Note=Inferred from the cDNA sequence of Ref.5;
 CC -1- TISSUE SPECIFICITY: Widely expressed at low levels. Detected in
 CC peritoneal blood leukocytes, lung, small intestine, spleen,
 CC thymus, colon and at lower levels in placenta, liver and kidney.
 CC Very low expression in skeletal muscle, heart and brain. Detected
 CC in the leukemia cell lines HL-60 and U937, but not in Jurkat T-
 CC cell lymphoma and Daudi Burkitt's lymphoma. Detected in the
 CC melanoma cell line WM35, but not in WM793. Not detected in HeLa
 CC cervical carcinoma cells and Molt 4 lymphocytic leukemia cells.
 CC MISCELLANEOUS: In breast tumorigenesis, methylation-mediated
 CC silencing may affect genes and proteins that act as positive
 CC mediators of cell death.
 CC -1- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 CARD domain.
 CC -1- CAUTION: Ref.5 sequence differs from that shown due to a
 CC frameshift in position 4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AB023416; BAA87339.2; -
 CC EMBL: AF184072; AAC01187.1; -
 CC EMBL: AF184073; AAC01188.1; -
 CC EMBL: AF255794; AAF99665.1; -

DR EMBL: AF310103; AAG30286.1; -
 DR EMBL: AF384665; AAG38650.1; -
 DR EMBL: AK000211; BAA91012.1; ALT_FRAME.
 DR EMBL: BC004470; AAB04470.1; -
 DR EMBL: BC013569; AAB13569.1; ALT_INIT.
 DR MIM: 606838; -
 DR GO: GO:0006917; P:induction of apoptosis; TAS.
 DR InterPro: IPR001315; CARD.
 DR InterPro: IPR004020; PAND_DAPIN_dom.
 DR Pfam: PF02758; PAND_DAPIN; 1.
 DR PROSITE: PS50209; CARD; 1.
 DR PROSITE: PS50824; DAPIN; 1.
 KM Apoptosis; Anti-oncogene; Alternative splicing.
 FT DOMAIN 1 31
 FT 107 195 DAPIN.
 FT VARSPLIC 26 85
 FT VARSPLIC 93 111
 FT VARSPLIC 195 AA; 21627 MW; 455987286586F46A CRC64;
 SQ SEQUENCE 195 AA; 21627 MW; 455987286586F46A CRC64;
 Query Match 100.0%; Score 990; DB 1; Length 195;
 Best local Similarity 100.0%; Pred. No. 4.8e-79;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDALIDALENTAEELKFKKLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
 DB 1 MGRARDALIDALENTAEELKFKKLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
 QY 61 LETYGAELTANTYLRDMLGLOEMAGOLQATHQSGGAPAGIOAPPOSAKPGHLFIQDHR 120
 DB 61 LETYGAELTANTYLRDMLGLOEMAGOLQATHQSGGAPAGIOAPPOSAKPGHLFIQDHR 120
 QY 121 ALIARVTNEMILDALYGVLTDEQYQAVRAEPTNSKKRKLFSFTPAWNTCKDILLQA 180
 DB 121 ALIARVTNEMILDALYGVLTDEQYQAVRAEPTNSKKRKLFSFTPAWNTCKDILLQA 180
 QY 121 ALIARVTNEMILDALYGVLTDEQYQAVRAEPTNSKKRKLFSFTPAWNTCKDILLQA 180
 DB 121 ALIARVTNEMILDALYGVLTDEQYQAVRAEPTNSKKRKLFSFTPAWNTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195
 RESULT 2
 ASC_MOUSE STANDARD; PRT; 193 AA.
 ID ASC_MOUSE
 AC Q9EPB4; Q9D2W9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Apoptosis-associated speck-like protein containing a CARD (mASC)
 DE (PYCARD).
 GN ASC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Thymus;
 RX MEDLINE=20580347; PubMed=11139337;
 RA Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;
 RT "Murine ortholog of ASC, a CARD-containing protein, self-associates
 RT and exhibits restricted distribution in developing mouse embryos.";
 RL Exp. Cell Res. 262:128-133(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Breast tumor;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "PyCard a PYD and CARD containing molecule.";
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBD databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pancreas, and Tongue;

RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Peeble G., Quackenbush J.,
 RA Schirml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guelinich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshaw-Borle A., Yoshida K., Hasegawa Y., Kawaji H., Kohlwski S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:665-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Matsuda K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
 RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamarate P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hallyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schultz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL
 CC -1- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
 activity is mediated predominantly through the activation of
 caspase 9 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
 a redistribution from the cytoplasm to the aggregates occurs.
 CC These aggregates are hollow, perinuclear spherical, ball-like
 structures (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed in small intestine, colon, thymus,
 spleen, brain, heart, skeletal muscle, kidney, lung and liver.
 CC -1- DEVELOPMENTAL STAGE: Strongly expressed at E9.5 day in the
 telencephalon, thalamic areas of the diencephalon, heart and
 liver.
 CC -1- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 CARD domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB032249; BAB16609.1; -
 DR EMBL; AF310104; AAG30287.1; -
 DR EMBL; AK009852; BAB26543.1; -
 DR EMBL; AK007742; BAB25229.1; -
 DR EMBL; AK018682; BAB31341.1; -

DR EMBL; BC008252; AAH08252.1; -
 DR MGD; MGI:1931465; Asc.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR04020; PAAD DAPIN_dom.
 DR Pfam; PF02758; PAAD DAPIN; 1.
 DR PROSITE; PSS0209; CARD; 1.
 DR PROSITE; PSS0824; DAPIN; 1.
 RT Apoptosis; Anti-oncogene.
 RT DOMAIN 1
 FT DOMAIN 105 193
 FT CONFLICT 159 159 K -> E (IN REF. 3).
 SQ SEQUENCE 193 AA; 21458 MW; 24A8A40194870B31 CRC64;
 Query Match 69.1%; Score 684; DB 1; Length 193;
 Best Local Similarity 71.8%; Pred. No. 1.8e-52;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
 Oy 1 MGRARDAIIDLAEIVLTAELKKKFKLTKLTVLRGVRIRPRGALLSMDALDTPDKVSFY 60
 Db 1 MGRARDAIIDLAEIVLTAELKKKFKLTKLTVLRGVRIRPRGALLSMDALDTPDKVSFY 60
 Oy 61 LETYGAELTANVLRDMGLQEMAGQLQATRGSGCAPAGIQAPPQSAKPGHPIIDHRA 120
 Db 61 LETYGAELTANVLRDMGLQEMAGQLQATRGSGCAPAGIQAPPQSAKPGHPIIDHRA 120
 Oy 121 ALIARVTNVMELDALYGVKLTDEQYQVAPAEPPNPKMKLTFEPANWTCDDLLQA 180
 Db 119 ALIARVTNVMELDALYGVKLTDEQYQVAPAEPPNPKMKLTFEPANWTCDDLLQA 178
 Oy 181 LRESQSYLVEDLERS 195
 Db 179 LKEIHPYLVMDLEQS 193
 RESULT 3
 NALI_HUMAN STANDARD; PRT; 1473 AA.
 ID NALI_HUMAN
 AC Q9C000; Q9BZ28; Q9BZ29; Q9HAY8; Q9URT4; Q9T2E0;
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DE 15-SEP-2003 (Rel. 42, last annotation update)
 DE NACHT-1, LRR- and PYD-containing protein 2 (Death effector filament-
 forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCB1_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., Distefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 proteins.";
 RT Cell Death Differ. 7:1273-1274(2000).
 RL [2]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martignon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 implicated in apoptosis and inflammation.";
 RT Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=ERYTHROLEUKEMIA;
 RX MEDLINE=21153743; PubMed=11076957;
 RA Hlaing T., Guo R.-F., Dilley K.A., Lousiea J.M., Morrish T.A.,
 RA Shi W.M., Vincenz C., Ward P.A.;
 RT "Molecular cloning and characterization of DEFCAP-1 and -S, two
 isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 proteins.";

RL J. Biol. Chem. 276:9230-9238(2001).
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORMS 1, 2, 3 AND 4), AND PROTEIN INTERACTION.
 RC TISSUE=T-cell;
 RX MEDLINE=1153744; Pubmed=1113115;
 RA Chu Z.-L., Plo F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
 "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
 c-dependent caspase activation and apoptosis.";
 RT J. Biol. Chem. 276:9239-9245(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=99246063; Pubmed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama W., Kikuno R., Hirosewa M.,
 Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE=Uterus;
 RA Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Able to form cytoplasmic structures termed death
 effector filaments. Enhances Apaf1 and cytochrome c-dependent
 activation of pro-caspase-9 and consecutive apoptosis. Seems to
 bind ATP.
 CC -1- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 and with Apaf1 in a cytochrome c-inducible way leading to the
 formation of an apoptosome. This interaction may be ATP-dependent.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=NAC beta, DEFCAP-1;
 CC IsoId=Q9C000-1; Sequence=Displayed;
 CC Name=2; Synonyms=NAC alpha, DEFCAP-S;
 CC IsoId=Q9C000-2; Sequence=VSP_004327;
 CC Name=3; Synonyms=NAC gamma;
 CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
 CC Name=4; Synonyms=NAC delta;
 CC IsoId=Q9C000-4; Sequence=VSP_004326;
 CC -1- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 expressed in peripheral blood leukocytes, chronic myelogenous
 leukemia cell line K-562, followed by thymus, spleen and heart.
 CC Also detected in lung, placenta, small intestine, colon, kidney,
 liver and muscle.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 1 CARD domain.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF298548; AAG15254.1; -
 DR EMBL: AF310105; AAG30288.1; -
 DR EMBL: AF229059; AAK00748.1; -
 DR EMBL: AF229060; AAK00749.1; -
 DR EMBL: AF229061; AAK00750.1; -
 DR EMBL: AF229062; AAK00751.1; -
 DR EMBL: AB023143; BAA76770.1; -
 DR EMBL: AL117470; CAB55945.1; -
 DR PIR: T17255; T17255.
 DR HSSP: P13489; 1A4Y.
 DR MIM: 60636; -
 DR GO: 0005622; C:intracellular; IC.

DR GO: 0016506; F:apoptosis activator activity; NAS.
 DR GO: 0008656; F:caspase activator activity; NAS.
 DR GO: 0019899; F:enzyme binding activity; IPI.
 DR GO: 0006919; F:caspase activation; NAS.
 DR GO: 0006917; P:induction of apoptosis; NAS.
 DR InterPro: IPR001315; CARD.
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007091; LRR_RNinh.
 DR InterPro: IPR007111; NACHT_NTPase.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF00560; LRR_2.
 DR Pfam: PF02758; PAAD_DAPIN_1.
 DR PRINTS: PR00364; DISEASERIST.
 DR PROSITE: PS50209; CARD_1.
 DR PROSITE: PS50824; DAPIN_1.
 DR PROSITE: PS50837; NACHT_1.
 DR Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 92 DAPIN.
 FT 328 637 NACHT.
 FT REPEAT 704 725 LRR 1.
 FT REPEAT 807 830 LRR 2.
 FT REPEAT 864 887 LRR 3.
 FT REPEAT 921 944 LRR 4.
 FT REPEAT 950 973 LRR 5.
 FT REPEAT 1199 1215 LRR 6.
 FT REPEAT 1216 1236 LRR 7.
 FT DOMAIN 1374 1463 CARD.
 FT NP_BIND 334 341 ATP (POTENTIAL).
 FT VARSPIC 958 967 Missing (in isoform 3 and isoform 4).
 FT VARSPIC 1262 1305 Missing (in isoform 2 and isoform 3).
 FT MUTAGEN 340 340 K->L: NO EFFECT.
 FT 155 155 K->S: NO EFFECT.
 FT 246 246 L->H (IN REF. 1).
 FT 782 782 T->S (IN REF. 1).
 FT 878 878 T->S (IN REF. 1).
 FT 995 995 T->I (IN REF. 1).
 FT 1119 1119 M->V (IN REF. 1).
 FT 1184 1184 M->V (IN REF. 1 AND 6).
 FT 1241 1241 V->L (IN REF. 1).
 FT 1366 1366 R->C (IN REF. 1).
 FT SEQUENCE 1473 AA; 165865 MW; 438FDDC845C2562D CRC64;
 Query Match 25.3%; Score 250; DB 1; Length 1473;
 Best Local Similarity 44.5%; Pred. No. 9,6e-14;
 Matches 65; Conservative 23; Mismatches 48; Indels 10; Gaps 4;
 QY 54 DKLVG-FYLETGAEITANV--LRDMGLOEMA-----GQLQAATHGSGAAPAGIQAPPOS 106
 DB 1317 DQLFSFVYGHGSGIRLVOKKDETLVWELVVRGDMPT--TLIPRIAVPSPL 1373
 QY 107 AKPGHLPIDRRALLAVTVNEMLDALYGVLTLDQYQVRAPEPPNSKRRKLFST 166
 DB 1374 DAPOLHFVDQYREQLIARVTSVEVLDTLHGQVLSQEQYERVLANTPSSQWRKLFSL 1433
 QY 167 PAMNWTCDLLLOALRESGSYLVEDI 192
 DB 1434 QSWDRCKDGLYQALKEPHLIMEL 1459
 RESULT 4
 ASC BRARE STANDARD; PRT; 203 AA.
 AC 0919NG;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Apoptosis-associated speck-like protein containing a CARD (PYCARD).
 GN ASC OR ASCI.

Matches 32; Conservative 17; Mismatches 25; Indels 4; Gaps 2;

QY 4 ARDAIDALENTAEELKKFKLKLSPVLRGEGYR-IPRGALLSMADALDTKLVSYFLE 62
 DB 15 ARELLIAALQDSQEQOLKFRHKLKRDAPLD---GRSIPWGRLEHSDAVLDVTKLIEFYAP 71

QY 63 TYGAELTANVLRMDGLQ 80
 DB 72 EPAVDVTRKTLKQADIRD 89

RESULT 6
 PYAS_MOUSE STANDARD; PRT; 843 AA.

AC Q91MS2; Q8KOL4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 5-like.
 GN PYPAF5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 [1]
 RN CONCEPTUAL TRANSLATION OF 1-287.
 RA Hinz U.;
 RL Unpublished observations (FEB-2003).
 RP SEQUENCE OF 288-843 FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Manisina K., Farnier A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ueda T.B., Tomihata S., Carninci P., Pange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abiraman R.D., Mallaby S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murthy D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smolius D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [3]
 RN IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
 RA Albrecht M., Dominguez F.S., Schreiber S., Langner T.;
 RT Identification of mammalian orthologs associates PYPAF5 with distinct
 functional roles.";
 RL FEBS Lett. 538:173-177(2003).
 CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappa-B (By similarity).
 CC -1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. (By similarity).
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
 CC -1- CAUTION: The N-terminus was extended using ESTs and genomic
 sequences, in analogy to ortholog sequences.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; BC013519; AAH13519.1; -
 CC EMBL; BC031139; AAH31139.1; ALT_INIT.
 CC MGD; MGI:2141990; PyPAF5.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR007091; LRR_RNinh.
 CC Pfam; PF00560; LRR; 1.
 CC PROSITE; PS00824; DAPIN; 1.
 CC PROSITE; PS00837; NACHT; 1.
 CC ATP-binding; leucine-rich repeat; Repeat.
 CC DOMAIN 11 102 DAPIN.
 CC FT DOMAIN 168 485 NACHT.
 CC FT DOMAIN 569 585 ASP/GLU-RICH.
 CC FT DOMAIN 654 661 POLY-LYS.
 CC FT REPEAT 434 459 LRR 1.
 CC FT REPEAT 609 632 LRR 2.
 CC FT REPEAT 811 834 LRR 3.
 CC NP_BIND 174 181 ATP (POTENTIAL).
 CC SEQID 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;

Query Match 13.3%; Score 132; DB 1; Length 843;
 Best Local Similarity 41.0%; Pred. No. 0.00089;
 Matches 32; Conservative 14; Mismatches 28; Indels 4; Gaps 2;

QY 4 ARDAIDALENTAEELKKFKLKLSPVLRGEGYR-IPRGALLSMADALDTKLVSYFLE 62
 DB 15 ARELLIAALQDSQEQOLKFRHKLKRDAPLD---GRSIPWGRLEHSDAVLDVTKLIEFYAP 71

QY 63 TYGAELTANVLRMDGLQ 80
 DB 72 EPAVDVTRKTLKQADIRD 89

RESULT 7
 PYAS_HUMAN STANDARD; PRT; 892 AA.

AC P59044;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 5.
 GN PYPAF5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22162427; PubMed=12019269;
 RA Lora J.M., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Britskin M., Distefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing APAF1-like protein that regulates
 activation of NF-kappa B and caspase-1-dependent cytokine
 processing.";
 RT J. Biol. Chem. 277:29874-29880(2002).
 [2]
 RN FUNCTION.
 RP MEDLINE=2275822; PubMed=12387869;
 RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
 RA Kelly R., Carlson A., Merriam S., Lora J.M., Britskin M.,
 RA Distefano P.S., Bertin J.;
 RT "Functional screening of five PYPAF family members identifies PYPAF5
 as a novel regulator of NF-kappa B and caspase-1.";
 RL FEBS Lett. 530:73-78(2002).
 CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappa-B.
 CC -1- SUBUNIT: Binds to ASC with its DAPIN domain.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
 CC much lower levels in T-cells.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.

CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL: AF479748; AAL87105.1; -
 CC InterPro: IPR007091; LRR_RNinh.
 CC InterPro: IPR003590; LRR_RNinh_sub.
 CC InterPro: IPR007111; NACHT_NTPase.
 CC InterPro: IPR004020; PAD_DAPIN_dom.
 CC Pfam: PF02758; PAD_DAPIN; 1.
 CC SMART: SM00368; LRR_RI; 3.
 CC PROSITE: PS50824; DAPIN; 1.
 CC PROSITE: PS50837; NACHT; 1.
 CC ATP-binding; Leucine-rich repeat; Repeat.
 CC DOMAIN 1
 CC FT 103 DAPIN.
 CC FT 196 513 NACHT.
 CC FT 604 614 POLY-GLU.
 CC FT 462 487 LRR 1.
 CC FT 727 747 LRR 2.
 CC FT 755 778 LRR 3.
 CC FT 811 834 LRR 4.
 CC FT 845 868 LRR 5.
 CC FT NP_BIND 202 209 ATP (POTENTIAL).
 CC SEQUENCE 892 AA; 98733 MW; 4AABD1FC76DE9D CRC64;
 Query Match 12.7%; Score 125.5; DB 1; Length 892;
 Best Local Similarity 28.4%; Pred. No. 0.0035;
 Matches 48; Conservative 25; Mismatches 75; Indels 21; Gaps 5;
 QY 4 ARDAIILALNLTBEELKFKLLSVLRGSGYR-IPRGALLISMDALDLTKVSPYLE 62
 DB 18 ARELLALAEELSSQQLRFRHKRDV--GPDGSIWGLERADAVDLAEQLAQFTGP 74
 QY 63 TYGAELTANVLRDGLQEMAGLOAATQSGAAPAGIQAPPOSNAKPGHFIPOH--- 118
 DB 75 EPALAEVAKTLKRDARDARVNAQLQERRIQRLGSLGSLTSLSEYKKYKREHVQLQHRVK 134
 QY 119 -RAALLARVTNVEMLDLVYKVLTDQYQVR-----AEPTNPSKRXK 161
 DB 135 ERNARSVYITK-----RFTKLILAPESAPAEALGPAPEREGRAR 176
 RESULT 8
 MEVU HUMAN STRAND: PRT; 781 AA.
 AC 015553;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Pyrin (Marennosin).
 GN MEVU OR MEV.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS FMF ILE-680; VAL-694 AND ALA-726.
 RC TISSUE=Leukocyte;
 RX MEDLINE=97433089; PubMed=9288758;
 RA Aksenitjevich I., Centola M., Deng Z., Sood R., Balow J.B. Jr.,
 RA Wood G., Zaks N., Mansfield E., Chen X., Eisenberg S., Vedula A.,
 RA Sheifan N., Kaden N., Pras E., Pras M., Kastner D.L., Blake T.,
 RA Bavevaite A.D., Robbins C., Kitzman D., Collins F.S., Liu P.P.,
 RA Chen X., Shohat M., Hamon M., Kahan T., Cercak A., Rotter J.I.,
 RA Fischel-Ghodeshan N., Richards N., Shelton D.A., Gumucio D.,
 RA Yokoyama Y., Mangelsdorf M., Orsborn A., Richards R.I., Ricke D.O.,

RA Buckingham J.M., Moyzis R.K., Deaven L.L., Doggett N.A.;
 RT "Ancient missense mutations in a new member of the ROR1 gene family
 RT are likely to cause familial Mediterranean fever.";
 RL Cell 90:797-807(1997).
 RN [2]
 RP SEQUENCE OF 305-754 FROM N.A., AND VARIANTS FMF.
 RX MEDLINE=97434208; PubMed=9288094;
 RA Benoit A., Clepet C., Dasilva C., Devaud C., Petit J.-L.,
 RA Caloustian C., Cruaud C., Samson D., Pulcini F., Weissbach J.,
 RA Hellig R., Notariola C., Domingo C., Rozenbaum M., Bencherit E.,
 RA Topaloglu R., Dewalle M., Dross C., Hadjari P., Dupont M.,
 RA Demaille J., Touitou I., Smacou N., Nedelec B., Mery J.-P.,
 RA Chaabouni H., Delpech M., Grateau G.;
 RT "A candidate gene for familial Mediterranean fever.";
 RL Nat. genet. 17:25-31(1997).
 RN [3]
 RP VARIANTS FMF, AND VARIANT GLN-202.
 RX MEDLINE=98334552; PubMed=9668175;
 RA Benoit A., da Silva C., Petit J.-L., Cruaud C., Caloustian C.,
 RA Castet V., Ahmed-Arab M., Dross C., Dupont M., Cattani D., Smacou N.,
 RA Dode C., Pechoux C., Nedelec B., Medaxian J., Rozenbaum M., Rosner I.,
 RA Delpech M., Grateau G., Demaille J., Weissbach J., Touitou I.;
 RT "Non-founder mutations in the MEVU gene establish this gene as the
 RT cause of familial Mediterranean fever (FMF).";
 RL Hum. Mol. Genet. 7:1317-1325(1998).
 RN [4]
 RP VARIANTS FMF ILE-680; ILE-681; ILE-694; VAL-694; MET-694 DEL AND
 RP ALA-726.
 RX MEDLINE=99149053; PubMed=1024914;
 RA Booth D.R., Gilmore J.D., Booth S.E., Peyes M.B., Hawkins P.N.;
 RT "Pyrin/marennosin mutations in familial Mediterranean fever.";
 RL QJM 91:603-606(1998).
 RN [5]
 RP VARIANTS FMF.
 RX MEDLINE=99192341; PubMed=10090880;
 RA Aksenitjevich I., Torosyan Y., Samuels J., Centola M., Pras E.,
 RA Chae J.J., Odoux C., Wood G., Azaro M.P., Palumbo G., Giustolisi R.,
 RA Pras M., Oster H., Kastner D.L.;
 RT "Mutation and haplotype studies of Familial Mediterranean Fever reveal
 RT new ancestral relationships and evidence for a high carrier frequency
 RT with reduced penetrance in the Ashkenazi Jewish population.";
 RL Am. J. Hum. Genet. 64:949-962(1999).
 RN [6]
 RP VARIANTS FMF GLN-148; SER-369; GLN-408; LEU-479; ILE-680; VAL-694;
 RP ALA-726 AND HIS-761.
 RX MEDLINE=99294585; PubMed=10364520;
 RA Cazeneuve C., Sarkisian T., Pechoux C., Derwichian M., Nedelec B.,
 RA Reinert P., Ayvazyan A., Kouyoumdjian J.-C., Ajrapetyan H.,
 RA Delpech M., Goossens M., Dode C., Grateau G., Amelien S.;
 RT "MEVU-Gene analysis in Armenian patients with Familial Mediterranean
 RT fever: diagnostic value and unfavorable renal prognosis of the M694V
 RT homozygous genotype-genetic and therapeutic implications.";
 RL Am. J. Hum. Genet. 65:88-97(1999).
 RN [7]
 RP VARIANTS FMF ILE-680; ILE-694; VAL-694 AND ALA-726.
 RX MEDLINE=99250763; PubMed=10234504;
 RA Shohat M., Nagal N., Shohat T., Chen X., Degan T., Mimouni A.,
 RA Danon Y., Lotan R., Ogur G., Sirin A., Schlezinger M., Halpern G.J.,
 RA Schwabe A., Kastner D., Rotter J.I., Fischel-Ghodeshan N.;
 RT "Phenotype-genotype correlation in familial Mediterranean fever:
 RT evidence for an association between Met694Val and amyloidosis.";
 RL Eur. J. Hum. Genet. 7:287-292(1999).
 RN [8]
 RP VARIANTS FMF GLN-148; ILE-680; ILE-694; VAL-694; ARG-695; ALA-726 AND
 RP HIS-761.
 RX MEDLINE=20081069; PubMed=10612841;
 RA Akar N., Mistrigolu M., Yalcinkaya F., Akar E., Cakar N., Turner N.,
 RA Akcalus M., Tasan H., Metzner Y.;
 RT "MEVU mutations in Turkish patients suffering from Familial
 RT Mediterranean Fever.";
 RL Hum. Mutat. 15:118-119(2000).
 RN [9]
 RP VARIANT GLN-148.

RX MEDLINE=20202844; PubMed=10737995;
 RA Ben-Cherif E., Lerer I., Malamud E., Domingo C., Abeliovich D.;
 RT "The E180Q mutation in the MEFV gene: is it a disease-causing mutation
 or a sequence variant?"
 RL Hum. Mutat. 15:385-386(2000).
 CC -1- FUNCTION: LIKELY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC DIFFERENTIATED GRANULOCYTES.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PERIPHERAL BLOOD
 CC LEUCOCYTES, PARTICULARLY IN MATURE GRANULOCYTES BUT NOT IN
 CC MONOCYTES AND LYMPHOCYTES. NOT EXPRESSED IN SPLEEN, THYMUS,
 CC PROSTATE, TESTIS, OVARY, SMALL INTESTINE, COLON, HEART, BRAIN,
 CC PLACENTA, LUNG, LIVER, MUSCLE, KIDNEY, PANCREAS.
 CC -1- DISEASE: Defects in MEFV are the cause of familial Mediterranean
 CC fever (FMF) (MIM:249100), an autosomal recessive inherited
 CC disorder characterized by recurrent episodic fever, serosal
 CC inflammation and pain in the abdomen, chest or joints. It is
 CC frequently complicated by amyloidosis, which leads to renal
 CC failure and can be prophylactically treated with colchicine. FMF
 CC primarily affects ethnic groups living around the Mediterranean
 CC basin: North-African Jews, Armenians, Arabs and Turks.
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.
 CC -1- DATABASE: NAME=Genedis; NOTE=FMF;
 CC WWW="http://life2.tau.ac.il/Genedis/Tables/FMF/fmf.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF018080; AAB70557.1; -;
 CC EMBL; Y14441; CAA74793.1; -;
 CC EMBL; AF003147; CAA05906.1; -;
 CC EMBL; AF111163; AAD26152.1; -;
 CC Genew; HGNC:6998; MEFV.
 CC MIM; 249100; -;
 CC DR GO; GO:0005875; C-microtubule associated complex; IDA.
 CC DR GO; GO:0005634; C:nucleus; IDA.
 CC DR GO; GO:0003779; F-actin binding activity; IDA.
 CC DR GO; GO:0008270; F-zinc ion binding activity; NAS.
 CC DR GO; GO:0006554; P:inflammatory response; IDA.
 CC DR InterPro; IPR001870; B302.
 CC DR InterPro; IPR004020; PAAD_DAPIN_dom.
 CC DR InterPro; IPR006574; PRY.
 CC DR InterPro; IPR003877; SPRY_receptor.
 CC DR InterPro; IPR000315; Znf_Box.
 CC DR Pfam; PF02758; PAAD_DAPIN; 1.
 CC DR Pfam; PF00622; SPRY; 1.
 CC DR Pfam; PF00643; zf-B_Box; 1.
 CC DR SMART; SM00336; BBOX; 1.
 CC DR SMART; SM00589; PRY; 1.
 CC DR SMART; SM00449; SPRY; 1.
 CC DR PROSITE; PS00824; DAPIN; 1.
 CC DR PROSITE; PS01119; ZF_BBOX; 1.
 CC DR Nuclear protein; zinc-finger; Polymorphism; Disease mutation.
 CC KW DOMAIN 1 92
 CC FT ZN FING 370 412
 CC FT DOMAIN 420 437
 CC FT VARIANT 148 148
 CC
 CC VARIANT 167 167
 CC FT 167
 CC FT VARIANT 202 202
 CC FT 202
 CC FT VARIANT 267 267
 CC FT 267
 CC FT VARIANT 369 369
 CC FT 369
 CC P -> S (in FMF; reduced penetrance among

FT FT Ashkenazi Jews; associated with Gln-148
 FT and Gln-408 in cis; could be a
 FT polymorphism).
 FT /FTid=VAR_009055.
 FT R -> Q (in FMF; associated with Gln-148
 FT and Ser-369 in cis; could be a
 FT polymorphism).
 FT /FTid=VAR_009056.
 FT F -> L (in FMF).
 FT /FTid=VAR_009057.
 FT M -> I (in FMF).
 FT /FTid=VAR_009058.
 FT T -> I (in FMF).
 FT /FTid=VAR_009059.
 FT Missing (in FMF).
 FT /FTid=VAR_009060.
 FT M -> I (in FMF).
 FT /FTid=VAR_009061.
 FT M -> V (in FMF; very common mutation
 FT particularly in North-African Jews; can
 FT be associated with amyloidosis
 FT development).
 FT /FTid=VAR_009062.
 FT Missing (in FMF).
 FT /FTid=VAR_009063.
 FT K -> R (in FMF; reduced penetrance among
 FT Ashkenazi Jews).
 FT /FTid=VAR_009064.
 FT V -> A (in FMF; common mutation; found in
 FT Iraqi and Ashkenazi Jews, Druze,
 FT Armenians).
 FT /FTid=VAR_009065.
 FT
 FT Query Match 12.4%; Score 122.5; DB 1; Length 781;
 FT Best Local Similarity 30.8%; Pred. No. 0.0055;
 FT Matches 33; Conservative 18; Mismatches 53; Indels 3; Gaps 1;
 FT
 FT QY 6 DAVIDALRENTBEUKRKRLKLSVPLREGYRIRRGALISDADLDKVSFYLETG 65
 FT DB 7 DHLTLEETLEVPYDEKRFKLTQNTSVQKSHSRIRSQIQRRPVMATLTVYGEYEA 66
 FT QY 66 AEITANVLREDMGLQEMAGOLQATHQSGAAPAGIQAPQSAKPG 112
 FT DB 67 VGLTQVLRATNGRLABELHQAALQSVSTQNGND---DSASSSL 110
 FT
 FT RESULT 9
 FT PYA3 HUMAN STANDARD; PRT; 980 AA.
 FT AC OSKX34;
 FT DT 28-FEB-2003 (Rel. 41, Created)
 FT DT 28-FEB-2003 (Rel. 41, Last sequence update)
 FT DT 28-FEB-2003 (Rel. 41, Last annotation update)
 FT DE PYRIN-containing APAF1-like protein 3.
 FT GN PYPAF3.
 FT OS Homo sapiens (Human)
 FT OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 FT OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 FT ON NCBI_Taxid=9606;
 FT RX SEQUENCE FROM N.A.
 FT RX MEDLINE=22162427; PubMed=12019769;
 FT RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 FT RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.;
 FT RT "PYPAF3, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing.";
 RT J. Biol. Chem. 277:29874-29880(2002).
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AF464765; AAL69963.1; -
 DR InterPro: IPR007091; LRR_RNinh.
 DR InterPro: IPR007111; NACHT_NTPase.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF02758; PAAD_DAPIN.1.
 DR PROSITE: PSS0824; DAPIN.1.
 DR PROSITE: PSS0837; NACHT.1.
 KM ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1
 FT 172 491 NACHT.
 FT REPEAT 614 638 LRR 1.
 FT REPEAT 674 697 LRR 2.
 FT REPEAT 760 784 LRR 3.
 FT REPEAT 788 810 LRR 4.
 FT REPEAT 817 840 LRR 5.
 FT REPEAT 845 868 LRR 6.
 FT REPEAT 874 897 LRR 7.
 FT REPEAT 902 928 LRR 8.
 FT REPEAT 933 957 LRR 9.
 FT NP_BIND 178 185 ATP (POTENTIAL).
 SQ SEQUENCE 980 AA; 111806 MW; 822AF2PD4338003D CRC64;

Query Match 10.1%; Score 99.5; DB 1; Length 980;
 Best Local Similarity 26.6%; Pred. No. 0.71; Indels 23; Gaps 5;
 Matches 45; Conservative 21; Mismatches 80; Indels 23; Gaps 5;

OY 12 LENVTAELKKFKLKLVSPLREGYRI PGALLSMDALDYLTKLVSYLETYGAEITAN 71
 DB 14 LEQINDELKSPKSLWAFPLEVDLQKTPWSEVHADGKLAELVNTSENNIRNATVN 73
 OY 72 VADMGQEN-----GQLQATHOGSGAAPGIGIAPPOSAPKGLH--FIDQH 118
 DB 74 ILSEMNTELCKAKAKEMEDGQVLEIDNELGDA-----EEDSELAKPEKSGWRNSME 128
 OY 119 RAALIRAVTVNEMLDLYGKVLTDQYQAVRAEP-TNPSKMRQLFSEFT 166
 DB 129 KQSLVWNTFWQGDIDNFHDVTLRQ---RPLPFLNPTPKLTYT 173

RESULT 10
 CAR8_HUMAN STANDARD; PRT; 431 AA.
 AC Q9Y2G2; Q96P82;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Caspase recruitment domain protein 8 (Apoptotic protein NDBP1) (DABAR)
 DE (CARD-inhibitor of NF-kappaB activating ligand) (CARDINAL) (TUCAN).
 GN CARD8 OR NDBP1 OR KIAA0935.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Brain;
 RA MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 6:63-70(1999).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21950691; PubMed=11821383;

RA Razmaza M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
 RA DiStefano P.S., Bertin J., Alnemri E.S., a new CARD family member that regulates caspase-1
 RT CARD-8 protein, a new CARD family member that regulates caspase-1
 RT activation and apoptosis.";
 RL J. Biol. Chem. 277:13952-13958(2002).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Zhang H.,
 RA "A novel apoptotic protein, NDBP1, containing CARD and BH3 domains";
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Guier C., Vito P.;
 RT "DABAR, a novel CARD-containing protein";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21570185; PubMed=11551959;
 RA Bouchier-Hayes L., Conroy H., Bgan H., Adrain C., Creagh E.M.,
 RA MacFarlane M., Martin S.J.;
 RT "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
 RT of multiple NF-kappa B activation pathways";
 RN J. Biol. Chem. 276:44069-44077(2001).
 RN [6]
 RN SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC TISSUE=Kidney;
 RA Guo J.H., Yu L.;
 RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RN CHARACTERIZATION.
 RX MEDLINE=21402909; PubMed=11408476;
 RA Patchan N., Matsuzawa H., Krawejcka M., Matsuzawa S.-I., Kim H.,
 RA Okada K., Torii S., Kltada S., Krawejcki S., Welsh K., Plo F.,
 RA Godzik A., Reed J.C.;
 RT "TUCAN, an antiapoptotic caspase-associated recruitment domain family
 RT protein overexpressed in cancer.";
 RL J. Biol. Chem. 276:32220-32229(2001).
 RN [8]
 RN CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
 RX MEDLINE=22062958; PubMed=12067710;
 RA Stillo R., Leonard A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
 RT "TUCAN/CARDINAL and DABAR participate in a common pathway for
 RT modulation of NF-kappaB activation.";
 RL FEBS Lett. 521:165-169(2002).
 CC -I- FUNCTION: Inhibits NF-kappaB activation. May participate in a
 CC regulatory mechanism that coordinates cellular responses
 CC controlled by NF-kappaB transcription factor. Involved in the
 CC negative regulation of caspase-1.
 CC -I- SUBUNIT: May form homodimers. Interacts with NEMO and DABAR. Binds
 CC to caspase-1, pseudo-ICR and ICEBERG.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=long;
 CC IsoId=Q9Y2G2-1; Sequence=Displayed;
 CC Name=short;
 CC IsoId=Q9Y2G2-2; Sequence=VSP_000782; VSP_000783;
 CC -I- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
 CC placenta. Lower expression in heart, kidney and liver. Also
 CC expressed in spleen, lymph node and bone marrow.
 CC -I- SIMILARITY: Contains 1 CARD domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AB023172; BAA7679.1; -
 DR EMBL: AF322184; AAG50014.1; -
 DR EMBL: AF331519; AAK01126.1; -

DR EMBL; AY026322; AK08982.1; -
 DR EMBL; AF405558; AA02427.1; -
 DR EMBL; AF511652; AA046959.1; -
 DR InterPro; IPR001315; CARD.
 DR Pfam; PF00619; CARD; 1.
 DR SMART; SM00114; CARD; 1.
 DR PROSITE; PSS0209; CARD; 1.
 KW Apoptosis; Nuclear protein; Alternative splicing.
 FT DOMAIN 340
 FT VARSPIC 282
 FT VARSPIC 287 431
 FT MTTAGEN 366 366
 FT CONFLICT 60 60
 FT CONFLICT 326 326
 FT CONFLICT 422 422
 SQ SEQUENCE 431 AA; 48932 MW; CB54D1308073286 CRC64;
 Query Match 10.0%; Score 99; DB 1; Length 431;
 Best Local Similarity 24.0%; Pred. No. 0.3;
 Matches 42; Conservative 32; Mismatches 73; Indels 28; Gaps 5;
 QY 21 KKKFKLLSVPLREGGRIRGALLSMALDLDKLVSYLETYGALTNVLRDMLGE 80
 DB 281 KKKFKLLSVPLREGGRIRGALLSMALDLDKLVSYLETYGALTNVLRDMLGE 333
 QY 81 MAGQLQATHQSGAPAGIQAPPOSAAKPGHFIIDHRAALATVNVEMLDALY-GK 139
 DB 334 VA-----ASAP-----PPFS-----GAATVKNHQLQARMDLGLVDDLDQNE 373
 QY 140 VLTDEQYQAVRAEPTNSKRRKLFSPANNWTKILLQALRESQYVLEDER 194
 DB 374 VLTNEKEVLEQEKTRQSKNEBALSMVEKKGDALDVLFRSISERDYLVSYLQ 428
 RESULT 11
 ID SITE_SALTY STANDARD; PRT; 143 AA.
 AC Q924N3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Silver-binding protein site precursor.
 GN Site.
 OS Salmonella typhimurium.
 OG Plasmid pMG101.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 OK NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 21-35, AND CHARACTERIZATION.
 RX MEDLINE=99128056; PubMed=9930866;
 RA Gupta A., Matsui K., Lo J.-F., Silver S.;
 RT "Molecular basis for resistance to silver cations in Salmonella";
 RL Nat. Med. 5:183-186(1999).
 CC -1- FUNCTION: COMPONENT OF THE SIL CATON-EFFLUX SYSTEM THAT CONFERS
 CC -1- RESISTANCE TO SILVER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- INDUCTION: By silver.
 CC -1- SIMILARITY: TO E. COLI POB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF067954; AA011743.1; -
 KM Metal-binding; Periplasmic; Signal; Plasmid.
 FT SIGNAL 1 20

FT CHAIN 21 143 SILVER-BINDING PROTEIN SITE.
 SQ SEQUENCE 143 AA; 15201 MW; 703B9BC0FBDDCBE CRC64;
 Query Match 9.2%; Score 91.5; DB 1; Length 143;
 Best Local Similarity 31.6%; Pred. No. 0.37; Indels 9; Gaps 4;
 Matches 31; Conservative 11; Mismatches 47;
 QY 64 YGAEITANVLRDMLGEAGQLQATHQ-GSGAPAPAGIQAPPOSAAKPGHFIIDHRAAL 122
 DB 12 FGLISSAMATEFVNIHERVNNQAAPAHQMSAAVPGIO-----GPAHMGMDGDEQAI 66
 QY 123 IARVTNEMLDALYKGLTDEQYQ-AVRAEPTNSK 158
 DB 67 IAEITNQSADA-HQKVESHQRMWSQTSVPTGSK 103
 RESULT 12
 ID RUVB_THETN STANDARD; PRT; 338 AA.
 AC Q8RAN2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Holiday junction DNA helicase ruvb.
 GN RUVB OR TRE1180.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 CC Thermoanaerobacteriaceae; Thermoanaerobacter.
 OK NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4 / JCM 11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome";
 RL Genome Res. 12:689-700(2002).
 CC -1- FUNCTION: The ruvb-ruvb complex in the presence of ATP renatures
 CC cruciform structure in supercoiled DNA with palindromic sequence,
 CC indicating that it may promote strand exchange reactions in
 CC homologous recombination. RuvbA is an helicase that mediates the
 CC Holliday junction migration by localized denaturation and
 CC reannealing (By similarity).
 CC -1- SUBUNIT: Forms a complex with ruvA (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RUVB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AE013080; AA024411.1; -
 DR HAMAP; MF_00016; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_centre.
 DR InterPro; IPR004605; RuvB.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRPFAM; TIGR00635; ruvb; 1.
 KM DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
 KW Complete proteome.
 FT NP BIND 59 66
 SQ SEQUENCE 338 AA; 38072 MW; 2B9562D8C8BD988D1 CRC64;
 Query Match 9.0%; Score 89; DB 1; Length 338;
 Best Local Similarity 24.1%; Pred. No. 1.7;
 Matches 39; Conservative 32; Mismatches 73; Indels 18; Gaps 5;
 QY 3 RARDALDLENLTAELEK---KFKLKLSTVPLREG---YGRIPGALLSMAL----- 50

Db 169 RSRFGVINKUDYYSVELEKTIIRGSANIINIGDEDAAFIARSRKSTPRIANRLKRRV 228
 QY 51 DLTDLKVSFYLETYGAEELTANVL -RDMQIOMAGOLQAATHQSGAAPAGIQAPPOSAA 108
 Db 229 DFEAVKNGNGVYDNTANIALINMLGVDEMGELEIDRKILIAIIEKFGGPGVIGDAIAYAG 288
 QY 109 KKGHLFIIDQHRALLIARVNVWELLALYKNTLTDEQYQAVR 150
 Db 289 EDSGDTIEDMEYEPYLM---QIGFLNRTPRGRVVTKLAQYOLK 326

RESULT 13
 M22_STRPY STANDARD; PRT; 372 AA.
 ID M22_STRPY STANDARD; PRT; 372 AA.
 AC P50469;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE M protein, serotype 2.2 precursor.
 GN EMDL2.2.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T2/44/RB4/119;
 RX MEDLINE=92104662; PubMed=1370269;
 RA Bessen D.E., Fischetti V.A.;
 RT "Nucleotide sequences of two adjacent M or M-like protein genes of group A streptococci: different RNA transcript levels and identification of a unique immunoglobulin A-binding protein.";
 RL Infect. Immun. 60:124-135(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fischetti V.A., Bessen D.E.;
 RT "Immunoglobulin A binding protein.";
 RL Patent number US5556944, 17-SEP-1996.
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO PHAGOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
 CC -1- SIMILARITY: TO OTHER M PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X61276; CAA3582.1; --
 CC EMBL, 126204; -; NOT_ANNOTATED_CDS.
 CC PIR, S23326; S23326.
 CC InterPro: IPR005877; Gpos_Y5IRK.
 CC InterPro: IPR001899; Gram_pos_anchor.
 CC InterPro: IPR006193; LPXG.
 CC InterPro: IPR003345; M_repeat.
 CC Pfam: PF00746; Gram_pos_anchor; 1.
 CC Pfam: PF02370; M; 8.
 CC Pfam: PF04650; Y5IRK_signal; 1.
 CC PRINTS: PR00015; GPOSANCHOR.
 CC TIGRAME: TIGR01167; LPXG_anchor; 1.
 CC TIGRAME: TIGR01168; Y5IRK_signal; 1.
 CC PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
 CC Variance: PhageOsis; Cell wall; Peptidoglycan-anchor; Repeat;
 KW Antigen; Coiled coil; Signal.
 FT SIGNAL 1 41 POTENTIAL.
 FT CHAIN 42 342 M PROTEIN, SEROTYPE 2.2.

FT PROPEP 343 372 REMOVED BY SORTASE (POTENTIAL).
 FT DOMAIN 131 244 3 X REPEATS, TYPE C.
 FT REPEAT 153 153 C-1.
 FT REPEAT 173 195 C-2.
 FT REPEAT 222 244 C-3.
 FT DOMAIN 305 338 GLY/PRO-RICH (CELL WALL-SPANNING) (BY SIMILARITY).
 FT SITE 339 343 LPXG SORTING SIGNAL (POTENTIAL).
 FT MOD_RES 342 342 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 SQ SEQUENCE 372 AA; 4149 MW; 88FD5D0920C95C74 CRC64;

Query Match
 Best Local Similarity 26.8%; Pred. No. 3.1; Length 372;
 Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

QY 12 LNLTAELKKERKLLSVPLREGYGRIPGALLSMDL-----DLTKLVSFYLETYGA 66
 Db 76 LKINAEKKNKKLEINENLNNYKQDG-----IDLKKEKEDLKTATTKENEIS 133
 QY 67 ELTANVL-RDMGLQEMAGOLQAATHQSGAAPAGIQAPQ--SAAPGL-HFIDQHRALL 122
 Db 134 EASRKGSLSRDLASRTAKKELEAKQKLEAKENKKTGNGVSEASRKGSLNDEASRAA- 192
 QY 123 IARVNVWELLALYKNTLTDEQ-----YQAVRAEPTNSKRRKLFSPANNWTCDDL 177
 Db 193 -----KKELEAKQKLETHQALBAKQKLEADYQVSETSRKLS-----RD-- 234
 QY 178 LQALRESQSYLVEDLERS 195
 Db 235 LEASREANKVTSLELQA 252

RESULT 14
 ID SYPB_CORGL STANDARD; PRT; 835 AA.
 AC 08N06;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
 DB (Phenylalanyl-tRNA ligase beta chain) (Phars).
 GN PHET OR CGL1390.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(phe) = AMP + diphosphate + L-phenylalanyl-tRNA(phe).
 CC -1- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain family. Subfamily 1.
 CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AP005278; BAB98783.1; ALT_INT.
 CC DR HAMAP; MF 00283; -; 1.
 CC InterPro: IPR005146; B3_4.

```

DR InterPro; IPR005147; B5.
DR InterPro; IPR005121; Fdx-AntiCB.
DR InterPro; IPR004532; phet bact.
DR InterPro; IPR002547; trna_bind.
DR Pfam; PF03483; B3_4; 1.
DR Pfam; PF03484; B5_1.
DR Pfam; PF03147; FDX-ACB; 1.
DR Pfam; PF01588; trna_bind; 1.
DR TIGRfam; TIGR00472; phet bact; 2.
DR PROSITE; PS50886; TRBD; 1.
KW Aminacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; RNA-binding; trna-binding;
KW Complete proteome.
FT DOMAIN 44 160 TRNA-BINDING.
FT METAL 472 472 MAGNESIUM (BY SIMILARITY).
FT METAL 478 478 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
SIMILARITY).
FT METAL 481 481 MAGNESIUM (BY SIMILARITY).
FT METAL 482 482 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 835 AA; 89435 MW; 2858C9A0673DA49F CRC64;

Query Match
Best Local Similarity 24.5%; Score 86.5; DB 1; Length 835;
Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7;

QY 29 SVPLRSGYGRIPRGALLSMDALDITDKLVSFYV---ETYGAEITAVVLDMDGLOENAGOU 85
DB 114 AISRARETYGMSAGMCSASELGLADKQNSGITLDPSTYG-EPGSDARQALGLEDDVFPVY 172
QY 86 QAAITHGSGAAPRGIOAPQPSAKPGIHPFDQRRALLARVNVNEMLDALYKVTIT--- 142
DB 173 NTPPDGYALSANGLTRELASAF--SLTFPD--PAIEPAVAGIEVKVPAVEGSLINVEL 227
QY 143 DEQYCAVRAEPTNPSKSKKLFSTPAWN---WTCKOLLQALR 182
DB 228 REETKAIR-----FGLRKVSGIDPAASEPFWMQRELMISGR 264

RESULT 15
CIS1_MOUSE STANDARD; PRT; 1033 AA.
AC OBR4E8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cold autoimmune syndrome 1 protein homolog (PYRIN-containing
APAF1-like protein 1) (Mast cell maturation inducible protein 1).
GN CIA31 OR PYPAF1 OR MMIG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshiba R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT "Identification of inducible genes during in vitro maturation of mouse
bone marrow-derived mast cells to connective tissue-type mast cells.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling (By similarity).
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -1- SIMILARITY: Contains 1 DAPIN domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF486632; AAL90874.1; -.
DR WGD; MG1:2653833; Clasi1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_R1; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; Repeat; Leucine-rich repeat.
FT DOMAIN 1 91 DAPIN.
FT METAL 216 532 NACHT.
FT REPEAT 737 760 LRR 1.
FT REPEAT 794 817 LRR 2.
FT REPEAT 851 874 LRR 3.
FT REPEAT 880 903 LRR 4.
FT REPEAT 908 931 LRR 5.
FT REPEAT 937 964 LRR 6.
FT REPEAT 965 988 LRR 7.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B12117 CRC64;

Query Match
Best Local Similarity 35.6%; Score 86; DB 1; Length 1033;
Matches 21; Conservative 9; Mismatches 29; Indels 0; Gaps 0;

QY 1 MGRADALIDALENTAELTKKFKLTVSPRSGYGRIPRGALLSMDALDITDKLVSF 59
DB 1 MTSVCKLAQVYEDLEDVLDKFKKHLDEYPRKGCIPRPGQMKADHLDTATMIDF 59

```

Search completed: January 29, 2004, 13:46:32
Job time : 5.00185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 10.8067 Seconds

(without alignment)
4656.416 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 1 MGRARDILDALENTABLE.....LLQALRESQSYLVEDLERS 195

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	736	74.3	195	6	Q8HXK9
2	673	68.0	193	11	Q8CHK8
3	275	27.8	89	4	Q8NFB8
4	267	27.0	89	4	Q8WXC3
5	142	14.3	750	11	Q9J325
6	142	14.3	767	11	Q9J326
7	104	10.5	404	13	Q9DDJ2
8	91.5	9.2	127	2	Q8VPY3
9	91.5	9.2	127	2	Q8VPW8
10	91.5	9.2	127	2	Q8VPX8
11	91.5	9.2	127	2	Q8VPF8
12	91.5	9.2	133	16	Q9J197
13	86.5	8.7	439	13	P74033
14	86.5	8.7	828	16	Q8NOM6
15	86	8.7	580	17	Q8TSR7
16	86	8.7	1033	11	Q8R4B8

17	85	8.6	1488	16	Q8Z7Z5	Q8Z7Z5 salmonella
18	85	8.6	1488	16	Q93557	Q93557 salmonella
19	84.5	8.5	894	10	Q8G5F4	Q8G5F4 oryza sativ
20	84.5	8.5	1241	3	Q9HEM2	Q9HEM2 neurospora
21	84	8.5	292	16	Q25976	Q25976 helicobacte
22	84	8.5	530	5	Q8T9R6	Q8T9R6 aedes aegypt
23	83.5	8.4	542	17	Q9HRC1	Q9HRC1 halobacteri
24	83.5	8.4	2556	12	Q9QH56	Q9QH56 gallid herp
25	83	8.4	469	16	Q8CVN9	Q8CVN9 escherichia
26	83	8.4	545	16	Q9CB12	Q9CB12 mycobacteri
27	83	8.4	877	2	Q45095	Q45095 bacillus ci
28	82.5	8.3	409	5	Q96984	Q96984 stylyomychia
29	82.5	8.3	522	2	Q8S210	Q8S210 vibrio chol
30	82.5	8.3	522	16	Q9K394	Q9K394 vibrio chol
31	82.5	8.3	567	3	Q8N1R9	Q8N1R9 cryptococcu
32	82	8.3	292	16	Q9ZJ11	Q9ZJ11 helicobacte
33	82	8.2	456	4	Q9BUC6	Q9BUC6 homo sapien
34	81	8.2	503	2	Q9XBW8	Q9XBW8 rhodobacter
35	81	8.2	631	4	Q8NDM4	Q8NDM4 homo sapien
36	81	8.2	631	4	Q969Z0	Q969Z0 homo sapien
37	81	8.2	792	16	Q8ZPS3	Q8ZPS3 salmonella
38	81	8.2	792	16	Q8Z6J0	Q8Z6J0 salmonella
39	81	8.2	896	2	Q9AN79	Q9AN79 bradyrhizob
40	81	8.2	9477	2	Q914X3	Q914X3 streptomyces
41	80.5	8.1	266	10	Q8LM81	Q8LM81 oryza sativ
42	80.5	8.1	718	16	Q9JYB6	Q9JYB6 neisseria m
43	80.5	8.1	725	16	Q9JYB7	Q9JYB7 neisseria m
44	80.5	8.1	727	16	Q8DE24	Q8DE24 vibrio vuln
45	80.5	8.1	972	4	Q8N612	Q8N612 homo sapien

ALIGNMENTS

RESULT 1

ID	Q8HXK9	PRELIMINARY;	PRT;	195 AA.
AC	Q8HXK9	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Apoptosis-associated speck-like protein containing a CARD.			
GN	BASC.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
NC	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Maumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,			
RA	Katsuyama T., Sagara J., Taniguchi S., Inohara N.,			
RA	Postlethwait J.H., Nunez G., Inohara N.,			
RT	"Casp-1: A zebrafish caspase activated by ASC oligomerization required			
RT	for pharyngeal Arch development.";			
RL	J. Biol. Chem. 0:0-0(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Maumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T.,			
RA	Nikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.,			
RT	"ASC, a novel 22-kDa protein, aggregates during apoptosis of human			
RT	promyelocytic leukemia HL-60 cells.";			
RL	J. Biol. Chem. 274:33835-33838(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Maumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.,			
RA	"Murine ortholog of ASC, a CARD-containing protein, self-associates,			
RT	and exhibits restricted distribution in developing mouse embryos.";			
RL	Exp. Cell Res. 262:128-133(2001).			
DR	EMBL; AB050006; BAC43753.1;			
SO	SEQUENCE 195 AA; 21917 MW; 7C9D4BD8DBA9A9E8 CRC64;			

Query Match 74.3%; Score 736; DB 6; Length 195;
 Best Local Similarity 75.4%; Pred. No. 1.7e-59;
 Matches 147; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 1 MGRARDATIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 DB 1 MGTCTRAIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 QY 61 LETYGAELTANYLRDMGLOEMAGOLQAATHQSGAAPAGIQAAPPQSAKRGHFIIDHRA 120
 DB 61 LEAYGAEELTANYLRDMGMOEVAEQLETMSKPRNVLAEVRDPLQKTKAKGHLFVQDHA 120
 QY 121 ALIARTYNTVEMLDALYKVLTDQYQAVAEPTNSKRRKLFSPFPAWMTCKDILLQA 180
 DB 121 ALIARTYTVDGVIDALYKVLTDQYQAVAEPTSSDKRRKLFSPFPAWMTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRDTQPYLVDDLEQS 195

RESULT 2

Q8CHK8 PRELIMINARY; PRT; 193 AA.
 AC 08CHK8;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Apoptosis-associated speck-like protein.
 GN RASC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Masumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,
 RA Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazlett P.,
 RA Postlethwait J.H., Nunez G., Inohara N.,
 RA "Caspase: A zebrafish caspase activated by ASC oligomerization required
 RT for pharyngeal Arch development.";
 RL J. Biol. Chem. 0:0-0(2002).
 DR EMBL; AB053165; BAC43754.1; -
 SQ SEQUENCE 193 AA; 21654 MW; F3B27B56086A17B CRC64;

Query Match 68.0%; Score 673; DB 11; Length 193;
 Best Local Similarity 69.7%; Pred. No. 9.7e-54;
 Matches 136; Conservative 20; Mismatches 37; Indels 2; Gaps 2;

QY 1 MGRARDATIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 DB 1 MGRARDATIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 QY 61 LETYGAELTANYLRDMGLOEMAGOLQAATHQSGAAPAGIQAAPPQSAKRGHFIIDHRA 120
 DB 61 LEAYGAEELTANYLRDMGLOEMAGOLQAATHQSGAAPAGIQAAPPQSAKRGHFIIDHRA 120
 QY 121 ALIARTYNTVEMLDALYKVLTDQYQAVAEPTNSKRRKLFSPFPAWMTCKDILLQA 180
 DB 121 ALIARTYTVDGVIDALYKVLTDQYQAVAEPTNSKRRKLFSPFPAWMTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LRDTQPYLVDDLEQS 193

RESULT 3

Q8NFP8 PRELIMINARY; PRT; 89 AA.
 AC 08NFP8;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Pyrin-domain containing protein.
 GN PYC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Martinon F., Hofmann K., Tschopp J.,
 RT "Pyc1 a novel regulator of the inflammasome";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF467809; AAN03745.1; -
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 SQ SEQUENCE 89 AA; 10035 MW; CEDFED672E506F56 CRC64;

Query Match 27.8%; Score 275; DB 4; Length 89;
 Best Local Similarity 65.9%; Pred. No. 1e-17;
 Matches 58; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 1 MGRARDATIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 DB 1 MGTCTRAIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 QY 61 LETYGAELTANYLRDMGLOEMAGOLQA 88
 DB 61 YEDYAAELVAVLRDMRMLLEAARLQRA 88

RESULT 4

Q8WXC3 PRELIMINARY; PRT; 89 AA.
 AC 08WXC3;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-UN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pyrin-only protein 1.
 GN POP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Berlin J.,
 RT "POP1: a pyrin-only protein that regulates inflammatory signaling.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF454669; AAL58439.1; -
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 SQ SEQUENCE 89 AA; 10107 MW; 4CDFED672DDDD98E CRC64;

Query Match 27.0%; Score 267; DB 4; Length 89;
 Best Local Similarity 64.8%; Pred. No. 5.6e-17;
 Matches 57; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 MGRARDATIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 DB 1 MGTCTRAIDALENTTABELKKFKKLLSVPLREGYGRIPRGALISMDALDLTDKLVSYF 60
 QY 61 LETYGAELTANYLRDMGLOEMAGOLQA 88
 DB 61 YEDYAAELVAVLRDMRMLLEAARLQRA 88

RESULT 5

Q9UJ25 PRELIMINARY; PRT; 750 AA.
 AC 09UJ25;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Pvrin (Marenostrein).
GN MEV.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RA SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP MEDLINE=20279845; PubMed=10818206;
RC STRAIN=Sprague-Dawley;
RX Chae J.U., Cencila M., Akeentjevich I., Dutra A., Tran M., Wood G.,
RA Nagaraaju K., Kingma D.W., Liu P.P., Kaestner D.L.,
RT "Isolation, genomic organization, and expression analysis of the mouse
RT and rat homologs of MEV, the gene for Familial Mediterranean fever.";
RL Mamm. Genome 11:428-435(2000).
CC -1- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN AND, TO A LESSER DEGREE IN
CC THE LUNG, NOT EXPRESSED IN THYMUS, TESTIS, OVARY, HEART, BRAIN,
CC LIVER, KIDNEY AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
CC EMBL: AF143410; AAF03767.1; -.
DR InterPro: IPR004020; PAAD DAPIN_dom.
DR InterPro: IPR000315; Znf Bbox.
DR Pfam: PF02758; PAAD DAPIN.1.
DR SMART: SM00336; ZF-B_box; 1.
DR SMART: SM00336; BBOX; 1.
DR PROSITE: PS50824; DAPIN; 1.
DR PROSITE: PS50119; ZF BBOX; 1.
KM Inflammatory response: Actin-binding; Metal-binding; Cytoskeleton;
KW Microtubules; Zinc-finger; Zinc.
FT DOMAIN 1 92 DAPIN.
FT ZN FING 440 482 B BOX-TYPE.
SQ SEQUENCE 750 AA; 83994 MW; C76D0F3E02711312 CRC64;

Query Match 14.3%; Score 142; DB 11; Length 750;
Best Local Similarity 33.7%; Pred. No. 0.00025;
Matches 35; Conservative 19; Mismatches 50; Indels 0; Gaps 0;

QY 6 DAVIDALENTAEELKKFKLISVPLREGYGRIPGALLSMALDITDKLVSYLETYG 65
DB 7 DHLNTLEELLPPDFEKFKEKQNTSLKSHSRIPSLVQMARPIKILTRLLITTYGEYA 66
QY 66 AELTANVLRDMGLQEMAGLOAATGSGAAPAGIOAPQSAKP 109
DB 67 VRLTLQILRATNQQLAEELKATGPHELTENRGGVSGVSAS 110

RESULT 6
Q9J26 PRELIMINARY; PRT; 767 AA.
AC Q9J26;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Pvrin (Marenostrein).
GN MEV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RA SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP MEDLINE=20279845; PubMed=10818206;
RX Chae J.U., Cencila M., Akeentjevich I., Dutra A., Tran M., Wood G.,

```

```

RA Nagaraaju K., Kingma D.W., Liu P.P., Kaestner D.L.;
RT "Isolation, genomic organization, and expression analysis of the mouse
RT and rat homologs of MEV, the gene for Familial Mediterranean fever.";
RL Mamm. Genome 11:428-435(2000).
CC -1- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN PERIPHERAL BLOOD
CC GRANULOCYTES. NOT EXPRESSED IN LYMPHOCYTES, THYMUS, TESTIS, OVARY,
CC HEART, BRAIN, LUNG, LIVER, KIDNEY AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
CC EMBL: AF143409; AAF03766.1; -.
DR MCD; MG1:1859396; MeV.
DR InterPro: IPR004020; PAAD DAPIN_dom.
DR InterPro: IPR000315; Znf Bbox.
DR Pfam: PF02758; PAAD DAPIN.1.
DR SMART: SM00336; ZF-B_box; 1.
DR SMART: SM00336; BBOX; 1.
DR PROSITE: PS50824; DAPIN; 1.
DR PROSITE: PS50119; ZF BBOX; 1.
KM Inflammatory response: Actin-binding; Metal-binding; Cytoskeleton;
KW Microtubules; Zinc-finger; Zinc.
FT DOMAIN 1 92 DAPIN.
FT ZN FING 439 481 B BOX-TYPE.
SQ SEQUENCE 767 AA; 86264 MW; FF102CB3FPD7C1EB CRC64;

Query Match 14.3%; Score 142; DB 11; Length 767;
Best Local Similarity 28.3%; Pred. No. 0.0003;
Matches 32; Conservative 26; Mismatches 47; Indels 8; Gaps 1;

QY 6 DAVIDALENTAEELKKFKLISVPLREGYGRIPGALLSMALDITDKLVSYLETYG 65
DB 7 DHLNTLEELLPPDFEKFKEKQNTSLKSHSRIPSLVQMARPIKILTRLLITTYGEYA 66
QY 66 AELTANVLRDMGLQEMAGLOAAT-----HSGGAAPAGIOAPQSAKP 110
DB 67 VRLTLQILRATNQQLAEELKATGPHELTENRGGVSGVSSENKAKSVKP 119

RESULT 7
Q9DDJ2 PRELIMINARY; PRT; 404 AA.
AC Q9DDJ2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Casp2.
GN Casp2 OR CASP2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RA SEQUENCE FROM N.A.
RP MEDLINE=20373792; PubMed=10917738;
RX Itohara N., Nunez G.,
RT "Gene with homology to mammalian apoptosis regulators identified in
RT zebrafish".
RL Cell Death Differ. 7:509-510(2000).
DR EMBL: AF127410; AAG45230.1; -.
DR HSSP; P23466; 1ICE.
DR ZFIN: ZDB-GENE-020812-1; caspb.
DR InterPro: IPR002398; 1ICE.
DR InterPro: IPR002138; 1ICE_p10.
DR InterPro: IPR001309; 1ICE_p20.
DR InterPro: IPR004020; PAAD DAPIN_dom.
DR Pfam: PF00655; 1ICE_p10; 1ICE_p20; 1.
DR Pfam: PF00656; 1ICE_p20; 1.

```

DR Pfam: PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PRO0376; ILIBENZYME.
 DR SMART; SM00115; CASc; 1.
 DR PROSITE; PS01122; CASPASE_CYS; 1.
 DR PROSITE; PS01131; CASPASE_HIS; 1.
 DR PROSITE; PS02027; CASPASE_P10; 1.
 DR PROSITE; PS02028; CASPASE_P20; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR SEQUENCE 404 AA; 46077 MW; 629BECFPA87D3B CRC64;

Query Match 10.5%; Score 104; DB 13; Length 404;
 Best Local Similarity 30.8%; Pred. No. 0.38; Indels 14; Gaps 4;
 Matches 36; Conservative 17; Mismatches 50;

QY 10 DALENTLAELKKKLTLSVPLREBYGRIPRGALISMDALDITDKVSEFYLETYGAEILT 69
 DB 10 DYLEDLVEALKEKF-TROLWIGVKPGVPEIPRGKLEKRODVDSMWQOQYSDAGT-IT 67
 QY 70 ANVLRLMGLOEMAGOL-----QAATHQSGAAPAGIQAPPSAAKPGHLFTIQ 117
 DB 68 VQTLRKIKONERAKRLLESMLKVQSGQENKONSEP---QPIQIISQPIQIISQ 121

RESULT 8

Q8VPY3 PRELIMINARY; PRT; 127 AA.

AC Q8VPY3; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SILE (Fragment).
 GN SILE.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21602765; PubMed=11739772; Silver S.;
 RA Gupta A., Phung L.T., Taylor D.E., Taylor D.E.,
 RT "Diversity of silver resistance genes in inch incompatibility group
 RT plasmids.";
 RL Microbiology 147:3393-3402(2001).
 DR EMBL; AY009377; AAL68934.1; -.
 KW Plasmid.
 FT NON TER 127
 SQ SEQUENCE 127 AA; 13438 MW; 67B3822C25BFC9A5 CRC64;

Query Match 9.2%; Score 91.5; DB 2; Length 127;
 Best Local Similarity 31.6%; Pred. No. 1.1;
 Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLOEMAGOLQAATHQ-GSGAAPAGIQAPPSAAKPGHLFTIDHRAAL 122
 DB 12 FGLISSAMATEVTNIHERVNNAAQAPAHQWQSAAPVGIQ-----GTAAPMAGMDQHEQAI 66
 QY 123 IARVTNVEWLLDALYGKVLTDQYO--AVRAEPTNPSK 158
 DB 67 IAHETWNGSADA-HQKVVESHQRMGSGQTVSPGTSK 103

RESULT 9

Q8VPW8 PRELIMINARY; PRT; 127 AA.

AC Q8VPW8; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SILE (Fragment).
 GN SILE.
 OS Salmonella oranienberg.
 OG Plasmid MIP235.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 NCBI_TaxID=28147;

QY 64 YGAELTANVLRDMGLOEMAGOLQAATHQ-GSGAAPAGIQAPPSAAKPGHLFTIDHRAAL 122
 DB 12 FGLISSAMATEVTNIHERVNNAAQAPAHQWQSAAPVGIQ-----GTAAPMAGMDQHEQAI 66
 QY 123 IARVTNVEWLLDALYGKVLTDQYO--AVRAEPTNPSK 158
 DB 67 IAHETWNGSADA-HQKVVESHQRMGSGQTVSPGTSK 103

Query Match 9.2%; Score 91.5; DB 2; Length 127;
 Best Local Similarity 31.6%; Pred. No. 1.1;
 Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 123 IARVTNVEWLLDALYGKVLTDQYO--AVRAEPTNPSK 158
 DB 67 IAHETWNGSADA-HQKVVESHQRMGSGQTVSPGTSK 103

RESULT 10

Q8VPX8 PRELIMINARY; PRT; 127 AA.

AC Q8VPX8; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SILE (Fragment).
 GN SILE.
 OS Salmonella enterica subsp. enterica serovar Ohio.
 OG Plasmid pMIP233.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 NCBI_TaxID=117541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21602765; PubMed=11739772; Silver S.;
 RA Gupta A., Phung L.T., Taylor D.E., Taylor D.E.,
 RT "Diversity of silver resistance genes in inch incompatibility group
 RT plasmids.";
 RL Microbiology 147:3393-3402(2001).
 DR EMBL; AY009382; AAL68941.1; -.
 KW Plasmid.
 FT NON TER 127
 SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;

Query Match 9.2%; Score 91.5; DB 2; Length 127;
 Best Local Similarity 31.6%; Pred. No. 1.1;
 Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLOEMAGOLQAATHQ-GSGAAPAGIQAPPSAAKPGHLFTIDHRAAL 122
 DB 12 FGLISSAMATEVTNIHERVNNAAQAPAHQWQSAAPVGIQ-----GTAAPMAGMDQHEQAI 66
 QY 123 IARVTNVEWLLDALYGKVLTDQYO--AVRAEPTNPSK 158
 DB 67 IAHETWNGSADA-HQKVVESHQRMGSGQTVSPGTSK 103

RESULT 11

Q8VPY8 PRELIMINARY; PRT; 127 AA.

AC Q8VPY8; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE SILE (Fragment).
CN SILE.
OS Serratia marcescens.
OC Plasmid PR476b.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.B., Silver S.;
RT "Diversity of silver resistance genes in *Inch* incompatibility group
RT plasmids.";
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009372; AAL68931.1; -.
KW Plasmid.
FT NON TER
SQ SEQUENCE 127 AA; 127 13456 MW; 67B3822C2313A365 CRC64;
Query Match 9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;
QY 64 YGAEITANVLRDMLQEMAGQLQATHTQ--GSGAPAGIQAPFQSAKRGHFIPTQRAAL 122
DB 12 FGLISSAMATEETVNIHERVNNAPAHQMGSAAPVGLQ-----GTAFRMAGMDHDEQAI 66
QY 123 IARVTNEMILDALYGVLTDEQYQ--AVRAEPTNPSK 158
DB 67 IAHETMNGSADA-HQKGVESHQRMGSGQTVPSPK 103
RESULT 12
Q919L7 PRELIMINARY; PRT; 383 AA.
AC 0919L7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Caspase.
OS CASPA OR CASPY.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF233434; AAF66964.1; -.
DR HSSP; P29466; 1ICE.
DR ZFIN; ZDB-GENE-000616-3; caspa.
DR InterPro; IPR002398; 1CE.
DR InterPro; IPR001309; 1CE.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00655; 1CE_P10; 1.
DR Pfam; PF00656; 1CE_P20; 1.
DR PRINTS; PR03758; PAAD_DAPIN; 1.
DR PRINTS; PR0376; IL1BENZYM.
DR SMART; SM00115; CASC; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS50207; CASPASE_P10; 1.
DR PROSITE; PS50208; CASPASE_P20; 1.
DR PROSITE; PS50824; DAPIN; 1.
SQ SEQUENCE 383 AA; 43966 MW; 21890871309774C3 CRC64;

Query Match 9.2%; Score 91.5; DB 13; Length 383;
Best Local Similarity 26.2%; Pred. No. 5;
Matches 28; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
QY 5 RDALIDALEMTAELKKFKLISVLRGCGYRIPGALISM-DALDITDKVSYLET 63
DB 6 KDHQDLSNIGADNLRFSRLD---RKQRPVRSSTIEKLDRIIDLIVNTFTSD 62
QY 64 YGAEITANVLRDMLQEMAGQLQATHTQSGAPAGIQAPFQSAK 110
DB 63 -AVSVTVDIRGKICNAVARELLENTGQ-----GVSGEPVPPEP 102
RESULT 13
P74033 PRELIMINARY; PRT; 439 AA.
ID P74033
AC P74033;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein slr0806.
DN SLR0806.
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kozaki H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kikuta T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nanno K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yaeuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90911; BAA18106.1; -.
DR InterPro; IPR006094; Oxid_FAD_bind.
DR Pfam; PF01565; FAD_binding_4; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 439 AA; 47989 MW; 6C2B6F56365FA653 CRC64;
Query Match 8.7%; Score 86.5; DB 16; Length 439;
Best Local Similarity 24.0%; Pred. No. 17;
Matches 42; Conservative 22; Mismatches 60; Indels 51; Gaps 8;
QY 31 PLRSGYR---IPRGALLSDALDLTDKVSFYLETGAEITANVLRDMLQEMAGQLQ 87
DB 194 PVHGYGTNGITETITPLTPALPWEAIVSF-----TNLSAIAFAQN 237
QY 88 ATHQ--GSGAPAGIQAPP-----OSAAKPGHFIPTQRAALIAVTNEM-----LL 133
DB 238 LAHQDGVSKSEISIQADPIQYFSLKSYQPGAHW-----MVISLSDWLAFTQLA 290
QY 134 DALYGVLTDEQYQAVRAEPTNPSKREKLSFTFAMWTCDDLLQALRBSQSTL 188
DB 291 KASGEIIFEQ-----DPOSFGKINLIEF-----WNHTTLARAVDPSTLYL 334
RESULT 14
Q8N0N6 PRELIMINARY; PRT; 828 AA.
ID Q8N0N6
AC Q8N0N6;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta subunit (Ec 6.1.1.20).
GN CG11390.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB005278; BMB98783.1; -
 DR InterPro: IPR005146; B3_4.
 DR InterPro: IPR005147; B5.
 DR InterPro: IPR005121; Fdx-AntiCB.
 DR InterPro: IPR004532; Phet bact.
 DR InterPro: IPR002547; trna_bind.
 DR Pfam: PF03483; B3_4; 1.
 DR Pfam: PF03484; B5; 1.
 DR Pfam: PF03147; FDX-ACB; 1.
 DR Pfam: PF01588; trna_bind; 1.
 DR TIGRfam: TIGR00472; phet bact; 2.
 DR Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
 KW SEQUENCE 828 AA; 88611 MW; E3F13B39CE5F29FE CRC64;

Query Match 8.7%; Score 86.5; DB 16; Length 828;
 Best Local Similarity 24.5%; Pred. No. 41;
 Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7;

QY 29 STPLRGYGRIRGALISMDALDTDKLYSFYL---ETYGAEITANVLDKMGIOEMAGOL 85
 DB 107 AISAERTYGRMSAGMICSASBELGLADKONSGLITLDPISYG-EPGEDARQALGLEDTVPDV 165
 QY 86 QATHGSGAAPAGIAPQSAKPGIHLFDHRAALIRVTVNEMVLALYGVKVT--- 142
 DB 166 NTPPDGKVALSARGLRELASAR--SLITTD--PAIEPAVACIEYKVAEBSGLINVEL 220
 QY 143 DEQYQVRAEPTNPSKRLFSFTPMN---WTCXDLILQALR 182
 DB 221 RETTKAIR-----FGLRKVSGIDPAESPFWMQRELMSGOR 257

RESULT 15

Q8TSR7 PRELIMINARY; PRT; 580 AA.
 AC Q8TSR7;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE DNA ligase (ATP).
 GN MA0728.
 OS Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;
 OC Methanosarcinaceae: Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZA / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=1932228;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smitrov S., Alnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thamm N., Dearellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talama J., Turrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayan L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of *Methanosarcina acetivorans* reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542 (2002).
 DR EMBL: AE010734; AAM04168.1; -
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR002016; Peroxidase.
 DR Pfam: PF01068; DNA_ligase; 1.
 DR Pfam: PF04679; DNA_ligase_A_C; 1.

DR Pfam: PF04675; DNA_ligase_A_N; 1.
 DR TIGRfam: TIGR00574; dn11; 1.
 DR PROSITE: PS00697; DNA_LIGASE_A1; 1.
 DR PROSITE: PS0160; DNA_LIGASE_A3; 1.
 DR PROSITE: PS00435; PEROXIDASE_1; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 580 AA; 65351 MW; 31C511F878C79E73 CRC64;

Query Match 8.7%; Score 86; DB 17; Length 580;
 Best Local Similarity 23.1%; Pred. No. 28;
 Matches 45; Conservative 30; Mismatches 60; Indels 60; Gaps 9;

QY 10 DALEMLTAEELKKFKLISVPLRGYGRIRGALISMDALDTDKLYSFYLETYGAEIT 69
 DB 136 DILQKATPEE-GKYLIRIVLGRRLGFGD-----QFLLEAFSIATF 175
 QY 70 A-----NVLRLMGLOEMAGOLQATHGSGAAPAGIAPQSAKPGIHLFD 117
 DB 176 GDKKAGKIKESYCTDIG--ELA---QTLAEHAG-----APGYSIRGPRV--- 220
 QY 118 HRAALIRVTVNEMVLALYGVKVTDEQYQVRAEPTNPSKRLFSFTPMNWTCKDL- 176
 DB 221 -KSMLAQRVESFELEERIKGKKAPEKVDGERVQ-----IHKXGDEIKAFSRLEDIT 273
 QY 177 -----LLOALRESQS 186
 DB 274 AQYPIIEAVRESIS 288

Search completed: January 29, 2004, 13:48:13
 Job time: 12.8067 secs


```

; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-61

```

Query Match 69.1%; Score 684; DB 4; Length 193;

Best Local Similarity 71.8%; Pred. No. 3.5e-75; Mismatches 34; Indels 2; Gaps 2;

```

Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDALDALENTLAEIKKKKLTLSVPLREGYGRIPRGALISMDALDTDKLVSY 60
DB 1 MGRARDALDALENTLAEIKKKKLTLSVPLREGYGRIPRGALISMDALDTDKLVSY 60
QY 61 LETYGAELTANTLRDMLGLOMAGOLAAHTGSGAAPAGIOAPPQSAKPGHFIIDORA 120
DB 61 LETYGAELTANTLRDMLGLOMAGOLAAHTGSGAAPAGIOAPPQSAKPGHFIIDORA 120
QY 61 LESYGLTMTVLRDMLGLOMAGOLAAHTGSGAAPAGIOAPPQSAKPGHFIIDORA 118
DB 61 LESYGLTMTVLRDMLGLOMAGOLAAHTGSGAAPAGIOAPPQSAKPGHFIIDORA 118
QY 121 ALIARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMTCKDILLOA 180
DB 121 ALIARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMTCKDILLOA 180
QY 119 ALIARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMTCKDILLOA 178
DB 119 ALIARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMTCKDILLOA 178
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195
QY 179 LKEIHPLYLVMDLEQS 193
DB 179 LKEIHPLYLVMDLEQS 193

```

RESULT 3

US-09-340-620A-58

```

; Sequence 58, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-58

```

Query Match 38.2%; Score 378; DB 4; Length 71;

Best Local Similarity 100.0%; Pred. No. 1.5e-38; Mismatches 0; Indels 0; Gaps 0;

```

QY 111 GHFIDQRAALIAARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMT 170
DB 1 GHFIDQRAALIAARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMT 60
QY 171 WTCKDILLOAL 181
DB 61 WTCKDILLOAL 71

```

RESULT 4

US-09-340-620A-57

```

; Sequence 57, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-57

```

Query Match 26.7%; Score 264; DB 4; Length 70;

Best Local Similarity 73.9%; Pred. No. 1.2e-24; Mismatches 11; Indels 0; Gaps 0;

```

Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 113 HFIDQRAALIAARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMT 172
DB 2 HFIDQRAALIAARTVNTVEMLDALYGKLTDEQYQAVRAEPTNSKKRKLFSFTPAWMT 61
QY 173 CKDILLOAL 181
DB 62 CKDILLOAL 70

```

RESULT 5

US-09-340-620A-66

```

; Sequence 66, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-66

```


US-09-340-620A-66

Query Match 26.7%; Score 264; DB 4; Length 70;

Best Local Similarity 73.9%; Pred. No. 1.2e-24; Indels 0; Gaps 0;

Matches 51; Conservative 7; Mismatches 11;

QY 113 HFDHRAALIAVTVNEMLDALYGVLTDEQYQAVRAEPTNPSKRRKLFSEFTPANWT 172

DB 2 HFDQHQALIAVTVNEMLDALYGVLTDEQYQAVRAEPTNPSKRRKLFSEFTPANWT 61

QY 173 CKDILLQAL 181

DB 62 CKDILLQAL 70

RESULT 6

US-09-340-620A-71

Sequence 71, Application US/09340620A

Patent No. 6482933

GENERAL INFORMATION:

APPLICANT: Bertin, John

TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE

FILE REFERENCE: 07334-124001

CURRENT APPLICATION NUMBER: US/09/340,620A

PRIOR FILING DATE: 1999-06-28

PRIOR APPLICATION NUMBER: US 09/245,281

PRIOR FILING DATE: 1999-02-05

PRIOR APPLICATION NUMBER: US 09/207,359

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: US 09/099,041

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: US 09/019,942

PRIOR FILING DATE: 1998-02-06

NUMBER OF SEQ ID NOS: 71

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 71

LENGTH: 109

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Consensus sequence

NAME/KEY: VARIANT

LOCATION: (1)...(109)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-340-620A-71

Query Match 15.3%; Score 151.5; DB 4; Length 109;

Best Local Similarity 46.7%; Pred. No. 1.2e-10; Indels 7; Gaps 4;

Matches 42; Conservative 13; Mismatches 28;

QY 111 GHHFDHRAALIAVTVNEMLDALYGVLTDEQYQAVRAEPTNPSKRRKLFSEFTP 167

DB 7 GSEIIDHRAALIAVTVNEMLDALYGVLTDEQYQAVRAEPTNPSKRRKLFSEFTP 66

QY 168 A-WNMTCKDL--LQALRESQSYLVEDLE 193

DB 67 SKGEETCKKFLKCLQALKQSAATLGLDPE 96

RESULT 7

US-09-328-352-7874

Sequence 7874, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 7874

LENGTH: 1233

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7874

Query Match 9.5%; Score 94.5; DB 4; Length 1233;

Best Local Similarity 26.5%; Pred. No. 0.047;

Matches 53; Conservative 26; Mismatches 76; Indels 45; Gaps 10;

QY 8 ILDALENTLBEELKFKLKLIVPLREGYGRIRGALLSDALDLVDKLVSYLETGAE 67

DB 125 VIDQLELFGVTLDSFQKLRFAFES-GKIER-AQITDA-----KITSRQ 170

QY 68 LITANVLVD-----MGLQEMAGOLQAA-----THQSGAAPAGIQAPQSAAPGL 112

DB 171 LHHVLRQWQSQPVYIDLVIAGELKSVDSFVKLVEDSLNFSNAHKLP-----EKPTI 226

QY 113 HFDHRAALIAVTVNEMLDALYGVLTDEQYQAVRAEPTNPSKRRKLFSEFTPANWT 171

DB 227 QF--EQLAQKQALATEIDISILEPPY--LIDGSHYKVNQTIFRNGAFNKLFS----- 275

QY 172 TCKDILLQALRESQSYLVED 191

DB 276 ECLPQLQILKQSDSILVFD 295

RESULT 8

US-07-813-584A-3

Sequence 3, Application US/07813584A

Patent No. 5352588

GENERAL INFORMATION:

APPLICANT: Fischetti, Vincent A.

TITLE OF INVENTION: No. 5352588el Immunoglobulin A Binding Protein

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kitley Murray

STREET: 98 Cutter Mill Road

CITY: Great Neck

STATE: NY

COUNTRY: USA

ZIP: 11021

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/813,584A

FILING DATE: 19911224

CLASSIFICATION: 435

AUTHOR/AGENT INFORMATION:

NAME: Murray, Kitley

REGISTRATION NUMBER: 30,246

REFERENCE/DOCKET NUMBER: RU-100

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-482-1990

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 372 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-813-584A-3

Query Match 8.7%; Score 86.5; DB 1; Length 372;

Best Local Similarity 26.8%; Pred. No. 0.068; Indels 37; Gaps 9;

Matches 53; Conservative 30; Mismatches 78;

QY 12 LENTLEELKFKLKLIVPLREGYGRIRGALLSDAL-----DLTDKLVSYLETGAE 66

DB 78 LKINAEERKKLLEINELNENYKLDG---IDALEKEKEDLTYYAKTKKEIS 133

QY 67 ELTANVL-RDMGLQEMAGOLQALATHQSGAAPAGIQAPQ--SAAKPGL-HFDHRAAL 122

```

Db      134 EASRKLRLDEASRTAKKLEAKHQKLEENKLTGNOVSEASRKLNDLEASRAA-192
QY      123 IARVTNVEMLLDALYKGVLTDEQ-----YQAVRAEPTNPSKMKLPSFTPAWNTCKDL 177
Db      193 -----KKELBAKYQKLETDHQLAKHQKLEADYGVSETSRKGLS-----RD--234
QY      178 LQALRESQSYLVEDLERS 195
Db      235 LEASREANKKVTSELTQA 252

RESULT 9
US-08-330-515-3
; Sequence 3, Application US/08330515
; Patent No. 5556944
; GENERAL INFORMATION:
; APPLICANT: Fieschetti, Vincent A.
; APPLICANT: Beesen, Debra E.
; TITLE OF INVENTION: No. 5556944el Immunoglobulin A Binding Protein
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,515
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,584
; FILING DATE: 24-DEC-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: RU-100.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-330-515-3

Query Match      8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 0.068;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

QY      12 LENTLAELKKPKLKLVLRSGYGRIRGALLSMAL-----DLTDKLVSYLVETGYA 66
Db      78 LKTNNAEERKQKLEALNKELNENYVTLQDG---IDALEKEKEDLTAKTKKEKIS 133
QY      67 ELTANVL-RDMGLQEMAGOLQAATHQSGGAAPAGIQAPQ--SAKKPGL-HFTDQRAAL 122
Db      134 EASRKLRLDEASRTAKKLEAKHQKLEENKLTGNOVSEASRKLNDLEASRAA-192
QY      123 IARVTNVEMLLDALYKGVLTDEQ-----YQAVRAEPTNPSKMKLPSFTPAWNTCKDL 177
Db      193 -----KKELBAKYQKLETDHQLAKHQKLEADYGVSETSRKGLS-----RD--234
QY      178 LQALRESQSYLVEDLERS 195

```

```

Db      235 LEASREANKKVTSELTQA 252

RESULT 10
US-09-069-023-5
; Sequence 5, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-069-023-5

Query Match      8.5%; Score 84.5; DB 4; Length 284;
Best Local Similarity 20.1%; Pred. No. 0.079;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;

QY      6 DAILEDLENTLAE-----IKKFKLKLVPLRSGYGRIRGALLSMDALDLDKLVSYLV 61
Db      35 EPLVLRFBETLLEAVITQKTKLQSVS-----SAIHLCDK-----70
QY      62 ETVGAEELTANVL-RDMGLQ-EMAGOLQAATHQSGGAAPAGIQAP-----103
Db      71 --KKMELSLNIPVNHGPQESGSSQLHNSGSPERSRLPAPQNDPLSRKAQDCYFMK 128
QY      104 -----PQSA-----KFGI--HFTDQR 119
Db      129 LHHCPGNHSDSTIGSGQPAFCDHKTTPCSSAIINPLSTAENSERLQPGIAQOWIQSR 188
QY      120 AALIAVNT--VEMLLDALYK-VLTDEQYQAVRAEPTNPSKMKLPSFTPAWNTCKDL 176
Db      189 EDIVNMTACLANQSIDLALSRDLIMKEDYELVSTKRTSKVRQLDITTDIOGEFAVY 248
QY      177 LQALRESQ 185
Db      249 IVQKLDKDK 257

RESULT 11
US-09-069-023-4
; Sequence 4, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-069-023-4

Query Match      8.5%; Score 84.5; DB 4; Length 478;
Best Local Similarity 20.1%; Pred. No. 0.18;

```


TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-942-1

Query Match 8.5%; Score 84.5; DB 3; Length 540;
Best Local Similarity 20.1%; Pred. No. 0.21;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;

QY 6 DAILEDLENTAEE---LKKFKLKLSVPLREGYGRIPGALLSMDALDITDKLSFYL 61
DB 291 EPLVLTFFETITFEAVIQLKTKLQSVS-----SAIHLCDK----- 326
QY 62 ETVGAEITANVLADMGLQ-EMAGQLQATHQSGGAAPAGIQAP----- 103
DB 327 --KMEELSLNIPVNHGPQESCGSLHENSISPETSRSLPAPQDNDFLSRKAQDCYFMK 384
QY 104 -----PQSA-----KPGI--HFIDQHR 119
DB 385 LHHCPGNHSDSTISGSQRAAFCDHKITPCSSAIINPLSTAGNSERLQPGIAQOWIOSKR 444
QY 120 AALIAAVTN--VEMLLDALYK-VLTDEQYQAVRAEPTNPSJKRKLFSEFTPANWNTCKDL 176
DB 445 EDIVNQMTACLNQSIDALLSRDLINKEDYELVSTKPTRISKVRQLDITDIQGEFPAY 504
QY 177 LQALRESQ 185
DB 505 IVQKLKDNK 513

RESULT 15
US-09-099-041A-2
Sequence 2, Application US/09099041A
Patent No. 6340576
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 540
TYPE: PRT
ORGANISM: Homo sapiens
US-09-099-041A-2

Query Match 8.5%; Score 84.5; DB 4; Length 540;
Best Local Similarity 20.1%; Pred. No. 0.21;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;

QY 6 DAILEDLENTAEE---LKKFKLKLSVPLREGYGRIPGALLSMDALDITDKLSFYL 61
DB 291 EPLVLTFFETITFEAVIQLKTKLQSVS-----SAIHLCDK----- 326
QY 62 ETVGAEITANVLADMGLQ-EMAGQLQATHQSGGAAPAGIQAP----- 103
DB 327 --KMEELSLNIPVNHGPQESCGSLHENSISPETSRSLPAPQDNDFLSRKAQDCYFMK 384
QY 104 -----PQSA-----KPGI--HFIDQHR 119
DB 385 LHHCPGNHSDSTISGSQRAAFCDHKITPCSSAIINPLSTAGNSERLQPGIAQOWIOSKR 444

QY 120 AALIAAVTN--VEMLLDALYK-VLTDEQYQAVRAEPTNPSJKRKLFSEFTPANWNTCKDL 176
DB 445 EDIVNQMTACLNQSIDALLSRDLINKEDYELVSTKPTRISKVRQLDITDIQGEFPAY 504
QY 177 LQALRESQ 185
DB 505 IVQKLKDNK 513

Search completed: January 29, 2004, 13:48:52
Job time : 3.84236 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 Search time 33.9809 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARDAIDALENTLAEEL.....LLIQALRESQSYVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA:
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	9 US-09-728-721-49	Sequence 49, Appl
2	990	100.0	195	10 US-09-996-617-8	Sequence 8, Appl
3	990	100.0	195	10 US-09-841-879B-5	Sequence 5, Appl
4	990	100.0	195	12 US-10-131-410-84	Sequence 84, Appl
5	990	100.0	195	12 US-10-240-145-55	Sequence 55, Appl
6	990	100.0	195	15 US-10-295-981-49	Sequence 49, Appl
7	990	100.0	205	15 US-10-106-698-5421	Sequence 5421, Ap
8	759	76.7	190	9 US-09-925-301-1120	Sequence 1120, Ap
9	684	69.1	193	9 US-09-728-721-61	Sequence 61, Appl
10	684	69.1	193	10 US-09-841-879B-2	Sequence 2, Appl
11	684	69.1	193	15 US-10-295-981-61	Sequence 61, Appl
12	564.5	57.0	158	15 US-10-106-698-5422	Sequence 5422, Ap
13	509	51.4	136	9 US-09-925-299-1421	Sequence 1421, Ap
14	509	51.4	136	11 US-09-925-299-1421	Sequence 1421, Ap
15	469	47.4	90	10 US-09-931-071-7	Sequence 7, Appl

16	445	44.9	85	10 US-09-841-879B-8	Sequence 8, Appl
17	445	43.9	89	14 US-10-127-516-16	Sequence 16, Appl
18	435	43.9	89	14 US-10-027-629-16	Sequence 16, Appl
19	378	38.2	71	9 US-09-728-721-58	Sequence 58, Appl
20	378	38.1	71	15 US-10-295-981-58	Sequence 58, Appl
21	377	38.1	77	14 US-10-127-516-8	Sequence 8, Appl
22	377	38.1	77	14 US-10-027-629-8	Sequence 8, Appl
23	370	37.4	77	12 US-10-132-967-8	Sequence 8, Appl
24	368	37.2	76	11 US-09-965-621-9	Sequence 9, Appl
25	368	37.2	76	12 US-10-407-866-9	Sequence 9, Appl
26	309	31.2	85	10 US-09-841-879B-15	Sequence 15, Appl
27	301	30.4	84	10 US-09-841-879B-7	Sequence 7, Appl
28	267	27.0	89	11 US-09-965-621-28	Sequence 28, Appl
29	267	27.0	89	12 US-10-407-866-28	Sequence 28, Appl
30	267	27.0	89	14 US-10-127-516-17	Sequence 17, Appl
31	267	27.0	89	14 US-10-027-629-17	Sequence 17, Appl
32	264	26.7	70	9 US-09-728-721-57	Sequence 57, Appl
33	264	26.7	70	9 US-09-728-721-66	Sequence 66, Appl
34	264	26.7	70	15 US-10-295-981-57	Sequence 57, Appl
35	264	26.7	70	15 US-10-295-981-66	Sequence 66, Appl
36	250.5	25.3	1399	10 US-09-388-221-4	Sequence 4, Appl
37	250.5	25.3	1429	10 US-09-996-617-2	Sequence 2, Appl
38	250.5	25.3	1429	10 US-09-931-071-2	Sequence 2, Appl
39	250.5	25.3	1429	12 US-10-028-374-15	Sequence 15, Appl
40	250.5	25.3	1429	12 US-10-183-770-15	Sequence 15, Appl
41	250.5	25.3	1429	15 US-10-028-392-11	Sequence 11, Appl
42	250	25.3	1443	10 US-09-388-221-6	Sequence 6, Appl
43	250	25.3	1473	10 US-09-388-221-2	Sequence 2, Appl
44	249.5	25.2	1429	12 US-10-028-374-3	Sequence 3, Appl
45	249.5	25.2	1429	12 US-10-183-770-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-728-721-49
Sequence 49, Application US/09728721
Patent No. US20020061845A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/728, 721
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: 09/340, 620
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/207, 359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099, 041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019, 942
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 195
TYPE: PRT
ORGANISM: Homo sapiens
US-09-728-721-49
Query Match 100.0%; Score 990; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAIDALENTLAEELKFKKLSVPLEGYSRIPRGALLSMDALDTDKLVSYF 60
DB 1 MGRARDAIDALENTLAEELKFKKLSVPLEGYSRIPRGALLSMDALDTDKLVSYF 60
QY 61 LETYGAELTNVIRDMGLQMGAGLOAATHTGSGAAPAGIOAPPOSAAKGLHFIDHRA 120
DB 61 LETYGAELTNVIRDMGLQMGAGLOAATHTGSGAAPAGIOAPPOSAAKGLHFIDHRA 120

QY 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 DB 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 2

US-09-996-617-8
 ; Sequence 8, Application US/09996617
 ; Patent No. US20020128198A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berlin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-340001
 ; CURRENT APPLICATION NUMBER: US/09/996,617
 ; PRIOR FILING DATE: 2001-11-27
 ; PRIOR APPLICATION NUMBER: 09/931,071
 ; PRIOR FILING DATE: 2001-08-15
 ; PRIOR APPLICATION NUMBER: 09/428,252
 ; PRIOR FILING DATE: 1999-10-27
 ; PRIOR APPLICATION NUMBER: 09/340,620
 ; PRIOR FILING DATE: 1999-06-28
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 195
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-996-617-8

Query Match 100.0%; Score 990; DB 10; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.1e-98;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 DB 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 QY 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 DB 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 QY 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 DB 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 3

US-09-841-879B-5
 ; Sequence 5, Application US/09841879B
 ; Patent No. US20020142979A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berlin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
 ; FILE REFERENCE: 07334-330001
 ; CURRENT APPLICATION NUMBER: US/09/841,879B
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: US 09/728,721
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: US 09/340,620
 ; PRIOR FILING DATE: 1999-06-28
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 195

QY TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-841-879B-5

Query Match 100.0%; Score 990; DB 10; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.1e-98;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 DB 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 QY 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 DB 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 QY 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 DB 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 4

US-10-131-410-84
 ; Sequence 84, Application US/10131410
 ; Publication No. US20030235915A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SPECHT, THOMAS
 ; APPLICANT: HINZMANN, BERND
 ; APPLICANT: SCHMITT, ARMIN
 ; APPLICANT: PILARSKY, CHRISTIAN
 ; APPLICANT: DAHL, EDGAR
 ; APPLICANT: ROSENTHAL, ANDRE
 ; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
 ; FILE REFERENCE: SCH-1763
 ; CURRENT APPLICATION NUMBER: US/10/131,410
 ; PRIOR FILING DATE: 2002-04-25
 ; PRIOR APPLICATION NUMBER: 09/646,673
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: PCT/DE99/00908
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 202
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 84
 ; LENGTH: 195
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-131-410-84

Query Match 100.0%; Score 990; DB 12; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.1e-98;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 DB 1 MGRARDALIDLALNTAEELKKFKKLSVPLREBYGRIIPRGALLSMOALDITDKLVSEFY 60
 QY 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 DB 61 LETYGAELTANVLRDMGQEWAGQLOAATHQSGGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
 QY 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 DB 121 ALIARVTNVEMLDLVYKVLTDQYQAVRAEPTNSKRRKLFSTPANNWTCCKLLLOA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

```
RESULT 5
US-10-240-145-55
; Sequence 55, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-09-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 55
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145-55

Query Match      100.0%; Score 990; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 3,1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGRARDAILDALLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 60
Db      1 MGRARDALIDLLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 60

QY      61 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
Db      61 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 120

QY      121 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 180
Db      121 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 180

QY      181 LRESQSYLVEDLERS 195
Db      181 LRESQSYLVEDLERS 195

RESULT 6
US-10-295-981-49
; Sequence 49, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07234-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-981-49

Query Match      100.0%; Score 990; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 3,1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGRARDAILDALLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 60
Db      1 MGRARDALIDLLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 60

QY      61 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
Db      61 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 120

QY      121 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 180
Db      121 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 180

QY      181 LRESQSYLVEDLERS 195
Db      181 LRESQSYLVEDLERS 195

RESULT 7
US-10-106-698-5421
; Sequence 5421, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5421
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5421

Query Match      100.0%; Score 990; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3,4e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGRARDAILDALLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 60
Db      1 MGRARDALIDLLENLTAEELKFKLKLISVLRREGYGRIPRGALLSMDALDITDKLVSFY 70

QY      61 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 120
Db      71 LETYGAEITANVLRDMGLOEMAGLOAATHTGSGAAPAGIOAPPOSAAKPGIHFIDQHRA 130

QY      121 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 180
Db      131 ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMKRLFSTPAMNWTCKDLLLOA 190

QY      181 LRESQSYLVEDLERS 195
Db      191 LRESQSYLVEDLERS 205
```

RESULT 8
US-09-925-301-1120
; Sequence 1120, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: FA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1120

Query Match 76.7%; Score 759; DB 9; Length 190;
Best Local Similarity 95.6%; Pred. No. 2,1e-73;
Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
DB 11 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 70
QY 61 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 120
DB 71 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 130
QY 121 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 160
DB 131 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 168

RESULT 9
US-09-728-721-61
; Sequence 61, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-728-721-61

Query Match 69.1%; Score 684; DB 9; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
DB 1 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60

QY 61 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 120
DB 61 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 118
QY 121 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 160
DB 119 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPIVLMDEQS 193

RESULT 10
US-09-841-879B-2
; Sequence 2, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-841-879B-2

Query Match 69.1%; Score 684; DB 10; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
DB 1 MGRARDALIDALENTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYF 60
QY 61 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 120
DB 61 LETYGAELTANTLRMGLOEMAGOLQAAHTGSGAAPGICQAPPOSAKPGHFIHQHRA 118
QY 121 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 160
DB 119 ALIARVTNVEWMLDALYGVLTDEGYQAVRAEPTNSKMR 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPIVLMDEQS 193

RESULT 11
US-10-295-981-61
; Sequence 61, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041

; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 61
 ; LENGTH: 193
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-295-981-61

Query Match 69.1%; Score 684; DB 15; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2,66-65;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 60
 DB 1 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 60
 QY 61 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIIDHRA 120
 DB 61 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIIDHRA 118
 QY 121 ALIARVTNVEWILDALYKVLTDQYQAVRAEPTNPGRKLFQFTDAMNTCKDLLOA 180
 DB 119 ALIARVTNVEWILDALYKVLTDQYQAVRAEPTNPGRKLFQFTDAMNTCKDLLOA 178
 QY 181 LRSQSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193

RESULT 12
 US-10-106-698-5422
 ; Sequence 5422, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005P1
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 5422
 ; LENGTH: 158
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-106-698-5422

Query Match 57.0%; Score 564.5; DB 15; Length 158;
 Best Local Similarity 75.8%; Pred. No. 1,4e-52;
 Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 ARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 63
 DB 13 ARDAIIDLALNTLAEELKFKK-----KLVSYFY 41
 QY 64 YGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIIDHRA 123
 DB 42 YGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIIDHRA 101
 QY 124 ARVTNVEWILDALYKVLTDQYQAVRAEPTNPGRK 160
 DB 102 ARVTNVEWILDALYKVLTDQYQAVR--PSPPTQAR 136

RESULT 13
 US-09-925-299-1421
 ; Sequence 1421, Application US/09925299
 ; Patent No. US20020055627A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA102
 ; CURRENT APPLICATION NUMBER: US/09/925,299
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05883
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1556
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1421
 ; LENGTH: 136
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-299-1421

Query Match 51.4%; Score 509; DB 9; Length 136;
 Best Local Similarity 83.7%; Pred. No. 1e-46;
 Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;
 QY 1 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 60
 DB 11 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 70
 QY 61 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIID 116
 DB 71 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIID 126
 QY 117 QHRAALIR 125
 DB 127 QHRAALSRR 135

RESULT 14
 US-09-925-299-1421
 ; Sequence 1421, Application US/09925299
 ; Publication No. US20030040617A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA102
 ; CURRENT APPLICATION NUMBER: US/09/925,299
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05883
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1556
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1421
 ; LENGTH: 136
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-299-1421

Query Match 51.4%; Score 509; DB 11; Length 136;
 Best Local Similarity 83.7%; Pred. No. 1e-46;
 Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;
 QY 1 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 60
 DB 11 MGRARDAIIDLALNTLAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSYFY 70
 QY 61 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIID 116
 DB 71 LETYGAELTANVLRDNGLOEMAGOLQAAATHOGSGAAPAGIQAPPOSAAKPGGLHFIID 126

QY 117 OHRAALIA 125
 Db 127 QHRASLSRR 135

RESULT 15
 US-09-931-071-7
 ; Sequence 7, Application US/09931071
 ; Patent No. US20020128219A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; APPLICANT: Alnemri, Emad S.
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-335001
 ; CURRENT APPLICATION NUMBER: US/09/931,071
 ; PRIOR FILING DATE: 2002-03-18
 ; PRIOR APPLICATION NUMBER: 09/428,252
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 90
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-931-071-7

Query Match 47.4%; Score 469; DB 10; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.2e-42;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SAKKPGHPTDQHRALIAVTVNVEWLLDALYGGKLTDEQYQAVRAEPTNPSXORLFSF 165
 Db 1 SAKKPGHPTDQHRALIAVTVNVEWLLDALYGGKLTDEQYQAVRAEPTNPSXORLFSF 60
 QY 166 TPANNTCKDLIQALRESOSYLVEDLERS 195
 Db 61 TPANNTCKDLIQALRESOSYLVEDLERS 90

Search completed: January 29, 2004, 13:54:38
 Job time : 34.9809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 19587.5 Seconds

(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-1

Perfect score: 5444
Sequence: 1 gccccagggccctggagaggtc.....aggataagaagttacctac 5444

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813366 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenBml:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_ov:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_man:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5444	100.0	5444	9 AB023143	AB023143 Homo sapi
2	5400.8	99.2	5404	9 AF310105	AF310105 Homo sapi
3	4882.4	89.7	5100	6 AX459865	AX459865 Sequence
4	4882.4	89.7	5100	9 AF298548	AF298548 Homo sapi
5	4765.6	87.5	5122	6 AX207661	AX207661 Sequence
6	4290	78.8	4290	9 AF229059	AF229059 Homo sapi
7	4281.8	78.7	5306	9 BC051787	BC051787 Homo sapi
8	4148	76.2	4422	6 AX089761	AX089761 Sequence
9	4148	76.2	4422	9 AF229060	AF229060 Homo sapi
10	4100	75.3	4200	9 AF229061	AF229061 Homo sapi
11	4078	74.9	4194	6 AX089763	AX089763 Sequence
12	3958	72.7	4332	9 AF229062	AF229062 Homo sapi
13	3952	72.6	4329	6 AX089765	AX089765 Sequence
14	3807.8	69.9	4123	9 HSM800983	AL117470 Homo sapi
15	3400.4	62.5	4556	6 AX089769	AX089769 Sequence
16	3210.4	59.0	4466	6 AX089771	AX089771 Sequence
17	3094.8	56.8	3419	9 AK026393	AK026393 Homo sapi
18	1710.8	31.4	152812	2 AC090536	AC090536 Homo sapi
19	1710.8	31.4	176730	9 AC055839	AC055839 Homo sapi
20	1566.4	28.8	2373	9 AK026398	AK026398 Homo sapi
21	991.8	18.2	11862	2 HSA001237	AC001237 Homo sapi
22	954	17.5	11862	2 HSA001237	AC001237 Homo sapi
23	742.8	13.6	153021	10 AL663042	AL663042 Mouse DNA
24	738	13.6	174564	2 AC121704	AC121704 Rattus no
25	738	13.6	250275	2 AC095695	AC095695 Rattus no
26	728	13.4	221540	2 AC026912	AC026912 Mus muscu
27	722	13.3	239196	2 AC127967	AC127967 Rattus no
28	701.6	12.9	96115	10 AL662908	AL662908 Mouse DNA
29	695.2	12.8	180274	2 AC027185	AC027185 Mus muscu
30	603	11.1	254690	2 AC095179	AC095179 Rattus no
31	399.8	7.3	3827	9 AY095146	AY095146 Homo sapi
32	399.8	7.3	3827	9 AY154467	AY154467 Homo sapi
33	399.8	7.3	69124	2 AC069567	AC069567 Mus muscu
34	398.2	7.3	3466	9 AK095460	AK095460 Homo sapi
35	398.2	7.3	3507	9 BC028069	BC028069 Homo sapi
36	398	7.3	3108	6 AX417214	AX417214 Sequence
37	398	7.3	3300	6 AX459869	AX459869 Sequence
38	398	7.3	3563	9 AY116205	AY116205 Homo sapi
39	398	7.3	3731	9 AY116204	AY116204 Homo sapi
40	397.4	7.3	3395	9 AY116206	AY116206 Homo sapi
41	395.8	7.3	3221	9 AY116207	AY116207 Homo sapi
42	392.2	7.2	108555	2 AC079221	AC079221 Mus muscu
43	376	6.9	3828	9 AF427617	AF427617 Homo sapi
44	374.4	6.9	3237	9 AY092033	AY092033 Homo sapi
45	373	6.9	4170	9 AF410477	AF410477 Homo sapi

ALIGNMENTS

RESULT 1

AB023143 5444 bp mRNA linear PRI 16-JUN-1999

LOCUS AB023143

DEFINITION Homo sapiens mRNA for KIAA0926 protein, complete cds.

AB023143

VERSION AB023143.1 GI:4589483

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (Saites)

Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirosewa,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.

Prediction of the coding sequences of unidentified human genes.

TITLE

XIII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 5444)
Ohara, O., Nagase, T. and Kikuno, R.
Submitted (04-FEB-1999) Osamu Ohara, Kazuo DNA Research Institute,
Laboratory of DNA Technology, Yama 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES
source
1.5444
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="h02962"
/sex="male"
/tissue_type="brain"
/clone_lib="Pbluescript SK plus"
/dev_stage="adult"
1.5444
/gene="KIAA0926"
523.4812
/gene="KIAA0926"
/codon_start=1
/product="KIAA0926 protein"
/protein_id="BAA76770.1"
/translation="589484"
/db_xref="GI:4589484"

gene
CDS

BASE COUNT
ORIGIN
1303 a 1525 c 1474 g 1142 t

Query Match
Best Local Similarity 100.0%; Score 5444; DB 9; Length 5444;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGAGGCTTGAAGAGTCTGAAGAACTTGGAGACAGAGCCCGGGGCTTCACTCT 60
DB 1 GCCCAGAGGCTTGAAGAGTCTGAAGAACTTGGAGACAGAGCCCGGGGCTTCACTCT 60
QY 61 GGGTTCGAAGGCTTCCCTGCTGCGGGCTCTCCACCCCACTCTTCAAGCTT 120
DB 61 GGGTTCGAAGGCTTCCCTGCTGCGGGCTCTCCACCCCACTCTTCAAGCTT 120
QY 121 GCAGCTCAAGGGTGTGATCTCAGAGAGTCCAGAGCCAGAGAGGAGAAATCTGAGAAACA 180
DB 121 GCAGCTCAAGGGTGTGATCTCAGAGAGTCCAGAGCCAGAGAGGAGAAATCTGAGAAACA 180

QY 181 CAGAACGTGAGGCTTCCCAACACCCCATCTCCCTACCAATCTCCCTCACTC 240
DB 181 CAGAACGTGAGGCTTCCCAACACCCCATCTCCCTACCAATCTCCCTCACTC 240
QY 241 CCTCCCTGCTGAGGCTTGAAGCCCATCCAGAGACCTCCCTACAGTACTCTTCACT 300
DB 241 CCTCCCTGCTGAGGCTTGAAGCCCATCCAGAGACCTCCCTACAGTACTCTTCACT 300
QY 301 GTCTTGCAGAGCCCTCTGAGGCTCTCTCCCTGAGCTTTTCTACACTCCCTCTAT 360
DB 301 GTCTTGCAGAGCCCTCTGAGGCTCTCTCCCTGAGCTTTTCTACACTCCCTCTAT 360
QY 361 CGGCGCTTATCTGTAGAGTCCCTGGAGTTTATATACTGGGTTCCGAATGCTGAATA 420
DB 361 CGGCGCTTATCTGTAGAGTCCCTGGAGTTTATATACTGGGTTCCGAATGCTGAATA 420
QY 421 GACGGTAAAGGCAAGGCAAGAAAGAGAGCACTGTCTCTGAGCTGAGCACTCCACAC 480
DB 421 GACGGTAAAGGCAAGGCAAGAAAGAGAGCACTGTCTCTGAGCTGAGCACTCCACAC 480
QY 481 CTGGGAACATCCCCAGACACCTCTTAACTCCGGAGACAGATGCTGCGAGGCTG 540
DB 481 CTGGGAACATCCCCAGACACCTCTTAACTCCGGAGACAGATGCTGCGAGGCTG 540
QY 541 GGGCGCTGAGCTTGTATCTTGAAGTTCTGAAAGAGAGAGCTGAAGAGTTCCAGCTT 600
DB 541 GGGCGCTGAGCTTGTATCTTGAAGTTCTGAAAGAGAGAGCTGAAGAGTTCCAGCTT 600
QY 601 CTGCTCGCAATTAAGGCGACCTCCAGAGACTCTTGGGGTGAAGACCCGCTCAAGCAG 660
DB 601 CTGCTCGCAATTAAGGCGACCTCCAGAGACTCTTGGGGTGAAGACCCGCTCAAGCAG 660
QY 661 AAGACAGATGAGATGAGAGTGGCTCTGATCTGATGCTGATGAGAGAGAGAGAGAG 720
DB 661 AAGACAGATGAGATGAGAGTGGCTCTGATCTGATGCTGATGAGAGAGAGAGAGAG 720
QY 721 TGGGACCTGAGCTTCTCATCTCTGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 780
DB 721 TGGGACCTGAGCTTCTCATCTCTGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CAGGAAGGAGGAGGAGGAGCTCTCCCTCACTCCCTCAAGGAGAGAGAGAGAGAG 840
DB 781 CAGGAAGGAGGAGGAGGAGCTCTCCCTCACTCCCTCAAGGAGAGAGAGAGAGAG 840
QY 841 TCTCCAGGCAACCCACTCTCACTCCAGAGAGTCTATGCTGATGATGATGATGATG 900
DB 841 TCTCCAGGCAACCCACTCTCACTCCAGAGAGTCTATGCTGATGATGATGATGATG 900
QY 901 GGGTGCACCCAGAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 GGGTGCACCCAGAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CGGTGAGAGAAATCTCTGCTCACTCTCTCACTCTCAAGAGCTCTTCAAGAGCTCC 1020
DB 961 CGGTGAGAGAAATCTCTGCTCACTCTCTCACTCTCTCAAGAGCTCTTCAAGAGCTCC 1020
QY 1021 GAGTCTCAAGGCAAGAGTCAACCAAGCCGCCCACTCAAGAGAGAGAGAGAGAG 1080
DB 1021 GAGTCTCAAGGCAAGAGTCAACCAAGCCGCCCACTCAAGAGAGAGAGAGAGAG 1080
QY 1081 GGAATCCCACTCAGAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 GGAATCCCACTCAGAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 CCTCTGAGAAACGTCAGAAATTTACTCAAGAAATCAGAAATCAGAAATCAGAAAT 1200
DB 1141 CCTCTGAGAAACGTCAGAAATTTACTCAAGAAATCAGAAATCAGAAATCAGAAAT 1200
QY 1201 TCAGAGAAAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 TCAGAGAAAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

QY 1261 AGCTACAGCCCAACCAACCACTGGAGGCTTCTGTGAGAGAGAGGCTCTGTTCACA 1320
DB 1261 AGCTTACAGCCCAACCAACCACTGGAGGCTTCTGTGAGAGAGAGGCTCTGTTCACA 1320
QY 1321 TGGCCCTGGAATAATGAGATTTTAAACCAAAATTCACAGCTGTCTACTTCAAAAGA 1380
DB 1321 TGGCCCTGGAATAATGAGATTTTAAACCAAAATTCACAGCTGTCTACTTCAAAAGA 1380
QY 1381 CCTTACCCCAAGAACCAAGATCCCTGTGTCAAGAGAAAGCTGGCTGTATTTGTGAGAG 1440
DB 1381 CCTTACCCCAAGAACCAAGATCCCTGTGTCAAGAGAAAGCTGGCTGTATTTGTGAGAG 1440
QY 1441 AATCGAGGACATTTAATGAGATCAGAGACTATTTTGGCCCAAGGCTGGATACCAAGAA 1500
DB 1441 AATCGAGGACATTTAATGAGATCAGAGACTATTTTGGCCCAAGGCTGGATACCAAGAA 1500
QY 1501 CCTCGATAGTCACTACTGACAGGGGCTGTGAAATTGGGAAGTCAACACTGGCCAGGAG 1560
DB 1501 CCTCGATAGTCACTACTGACAGGGGCTGTGAAATTGGGAAGTCAACACTGGCCAGGAG 1560
QY 1561 GTGAAGAGAGCTGGGGGAGAGAGGCAAGCTGTATGGGAGCCGCTTCCAGCATGTCTTAC 1620
DB 1561 GTGAAGAGAGCTGGGGGAGAGAGGCAAGCTGTATGGGAGCCGCTTCCAGCATGTCTTAC 1620
QY 1621 TTGAGCTGAGAGAGGCTGGCCAGTCCAGGTGTGTCTGTGCTGAGCTCATTCGAAAA 1680
DB 1621 TTGAGCTGAGAGAGGCTGGCCAGTCCAGGTGTGTGTCTGTGCTGAGCTCATTCGAAAA 1680
QY 1681 GATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCTAGGACAGAGGCTGCTC 1740
DB 1681 GATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCTAGGACAGAGGCTGCTC 1740
QY 1741 TTGATCTCGATGT 1800
DB 1741 TTGATCTCGATGT 1800
QY 1801 TGTCTGACCTGAGAGCCAGCCAGCCGCGGAGATGCACTGCTGGGCAAGTTTGTGGGAAA 1860
DB 1801 TGTCTGACCTGAGAGCCAGCCAGCCGCGGAGATGCACTGCTGGGCAAGTTTGTGGGAAA 1860
QY 1861 ACTATATCTCCAGGAGCATCTTCTGTATCAAGCTCTGGAACAAGCTCTGGAACCTC 1920
DB 1861 ACTATATCTCCAGGAGCATCTTCTGTATCAAGCTCTGGAACAAGCTCTGGAACCTC 1920
QY 1921 ATTCTCTTGTGAGAGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
DB 1921 ATTCTCTTGTGAGAGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
QY 1981 AAGGAATATTTCTACAGATATTTTCAAGATGAAGGAAGCAAGATTAGAGCTTTAGGTTG 2040
DB 1981 AAGGAATATTTCTACAGATATTTTCAAGATGAAGGAAGCAAGATTAGAGCTTTAGGTTG 2040
QY 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
DB 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
QY 2101 TGCATCTGCTGATGACAGATGAAGCGGAAGAAAACTCACACTGACTTCCAAAGAC 2160
DB 2101 TGCATCTGCTGATGACAGATGAAGCGGAAGAAAACTCACACTGACTTCCAAAGAC 2160
QY 2161 ACCAACAACCTGTGTACATTTACCTTGGCCAGGCTCCAAAGCTCAGCAATGGAGACC 2220
DB 2161 ACCAACAACCTGTGTGTACATTTACCTTGGCCAGGCTCCAAAGCTCAGCAATGGAGACC 2220
QY 2221 CAGCTCAGAGAGCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
DB 2221 CAGCTCAGAGAGCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
QY 2281 AGTTCAGATGACCTCAGAGAGAGATGGGTTAGATGGGCGCATCATCTTCAAG 2340
DB 2281 AGTTCAGATGACCTCAGAGAGAGATGGGTTAGATGGGCGCATCATCTTCAAG 2340
QY 2341 ATGGGTATTTCTTAAGAGACCCCATCCCTGTAGCTACAGCTTCATTCACCTGTGTTT 2400

DB 2341 ATGGGTATTTCTTAAGAGACCCCATCCCTGTAGCTACAGCTTCATTCACCTGTGTTT 2400
QY 2401 CAAAGATTCTTTGACAGCAATGTCTTATGTCTTGGAGATGAGAGGGAGAGGTAACAT 2460
DB 2401 CAAAGATTCTTTGACAGCAATGTCTTATGTCTTGGAGATGAGAGGGAGAGGTAACAT 2460
QY 2461 TCTAATTTGATCATTAATTTGAAAAAGCGCTTGAAGCATATGGATATCATGTGCTGTT 2520
DB 2461 TCTAATTTGATCATTAATTTGAAAAAGCGCTTGAAGCATATGGATATCATGTGCTGTT 2520
QY 2521 GGGGACATCAACCAACAGTTTCCATATTGGGCTGTATGATGATGAGGGGAGAGAGATG 2580
DB 2521 GGGGACATCAACCAACAGTTTCCATATTGGGCTGTATGATGATGAGGGGAGAGAGATG 2580
QY 2581 GAGAACATCTTTCACTGCGGCTGTCTCAGGGAGAGAACTGATGAGATGGGATCCGCTCC 2640
DB 2581 GAGAACATCTTTCACTGCGGCTGTCTCAGGGAGAGAACTGATGAGATGGGATCCGCTCC 2640
QY 2641 CTGACAGTGTGTGTGAGCACTCTCTGTGAGATCCCTCACTGTCTTGTACAGACTCGG 2700
DB 2641 CTGACAGTGTGTGTGAGCACTCTCTGTGAGATCCCTCACTGTCTTGTACAGACTCGG 2700
QY 2701 AACAAACGTTCTGACACAAAGTATGGCCATTTGAAAGAAATGGGCAATGTGTAGAA 2760
DB 2701 AACAAACGTTCTGACACAAAGTATGGCCATTTGAAAGAAATGGGCAATGTGTAGAA 2760
QY 2761 ACAGACATGAGAGCTTTAGT 2820
DB 2761 ACAGACATGAGAGCTTTAGT 2820
QY 2821 CTTCAGCTGATTTGAGGAGAGGAGAGCAAGATCAACATGAGAGCCCAACATGTGTGTGT 2880
DB 2821 CTTCAGCTGATTTGAGGAGAGGAGAGCAAGATCAACATGAGAGCCCAACATGTGTGTGT 2880
QY 2881 TTGAGGTGGTCCAGTGTCAAGATGCTTATTTGCAAGATTTCTTCCGCTTCAAGGTC 2940
DB 2881 TTGAGGTGGTCCAGTGTCAAGATGCTTATTTGCAAGATTTCTTCCGCTTCAAGGTC 2940
QY 2941 ACCAAGAACTGAAAGAGCTGAGCTTAAGTGAAGAACTGCTGAGCACTGTGAGTGAAG 3000
DB 2941 ACCAAGAACTGAAAGAGCTGAGCTTAAGTGAAGAACTGCTGAGCACTGTGAGTGAAG 3000
QY 3001 AGCTTTGTAAAGCCCTGAGAGCGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
DB 3001 AGCTTTGTAAAGCCCTGAGAGCGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
QY 3061 TGTGAGCTCAACAGTGAAGCTGCAAGAGACTTGGCTTTGGGCTGAGAGCCAAACAGACC 3120
DB 3061 TGTGAGCTCAACAGTGAAGCTGCAAGAGACTTGGCTTTGGGCTGAGAGCCAAACAGACC 3120
QY 3121 CTGACGAGCTGAGCTTGAAGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
DB 3121 CTGACGAGCTGAGCTTGAAGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
QY 3181 CAGAGACTGAGAGAGCGGAGCTGCAAGCTACAGCTGAGCTGTGTGTGTGTGTGTGTGT 3240
DB 3181 CAGAGACTGAGAGAGCGGAGCTGCAAGCTACAGCTGAGCTGTGTGTGTGTGTGTGTGT 3240
QY 3241 AGCTGTGACTGTGAGAGAGCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
DB 3241 AGCTGTGACTGTGAGAGAGCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
QY 3301 CTGACCTGAGAGAGAACTGT 3360
DB 3301 CTGACCTGAGAGAGAACTGT 3360
QY 3361 AGGCAATCTGAGAGAACTGT 3420
DB 3361 AGGCAATCTGAGAGAACTGT 3420
QY 3421 ATGAGGAGAGAGCTGAGAGGCTGTGAGAGAGAACTTCAAGCTGTGTGTGTGTGTGTGT 3480

Db 3421 ATGAGCAGGAACTGAGGCCCTTGAGACAGAGAACTCAGCTGCTCATCTTCAGACA 3480
 Qy 3481 CGGAAACCAAGTGTGATGACCCCTCACTGAGGGCTGTGATACGGAGAGATGATATAGC 3540
 Db 3481 CGGAAACCAAGTGTGATGACCCCTCACTGAGGGCTGTGATACGGAGAGATGATATAGC 3540
 Qy 3541 ACATCTCACTCAAGCGGACAGACTGTGATCAGAGAGAGCGGCTTCCATGTTGCTCAG 3600
 Db 3541 ACATCTCACTCAAGCGGACAGACTGTGATCAGAGAGAGCGGCTTCCATGTTGCTCAG 3600
 Qy 3601 GCTAATCTCAAACTCTGAGAGTGAAGATCTTCCAAATGCTGAGATGACAGAGAA 3660
 Db 3601 GCTAATCTCAAACTCTGAGAGTGAAGATCTTCCAAATGCTGAGATGACAGAGAA 3660
 Qy 3661 AGCTCCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3720
 Db 3661 AGCTCCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3720
 Qy 3721 GACCTGATAGCAAGCTTTGGGAGCTGAGATGATCTTGGGAGCTTCCACGGGAGCTGTG 3780
 Db 3721 GACCTGATAGCAAGCTTTGGGAGCTGAGATGATCTTGGGAGCTTCCACGGGAGCTGTG 3780
 Qy 3781 GCTACTGAGGTAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3840
 Db 3781 GCTACTGAGGTAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3840
 Qy 3841 TCCCTACGCTGGGCCCAACAGGGCTCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAG 3900
 Db 3841 TCCCTACGCTGGGCCCAACAGGGCTCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAG 3900
 Qy 3901 ATTGAATCTGT 3960
 Db 3901 ATTGAATCTGT 3960
 Qy 3961 GTGGCAGAGGCTCTGCTGAGCATCAAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 4020
 Db 3961 GTGGCAGAGGCTCTGCTGAGCATCAAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 4020
 Qy 4021 CCTCACTTGT 4080
 Db 4021 CCTCACTTGT 4080
 Qy 4081 TTTAAAG 4140
 Db 4081 TTTAAAG 4140
 Qy 4141 CTGGAAG 4200
 Db 4141 CTGGAAG 4200
 Qy 4201 CGCTTCATTCCTGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
 Db 4201 CGCTTCATTCCTGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
 Qy 4261 ACCTTCACCTCACTGATCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4320
 Db 4261 ACCTTCACCTCACTGATCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4320
 Qy 4321 TATCGAAGCCTGAG 4380
 Db 4321 TATCGAAGCCTGAG 4380
 Qy 4381 GGGATCAGGCTGCAAGTGAAG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAG 4440
 Qy 4441 AAACAGAGAGATCTCATGCTGTGCACTACTGTGATCTTCCACGGGAGAGAGAGAGAG 4500
 Db 4441 AAACAGAGAGATCTCATGCTGTGCACTACTGTGATCTTCCACGGGAGAGAGAGAGAG 4500
 Qy 4501 TCACCTCTGAGATGCGCGAGAGTGTGCTCACTTGTGTGACCAAGATGAGAGAGAGAGAT 4560
 Db 4501 TCACCTCTGAGATGCGCGAGAGTGTGCTCACTTGTGTGACCAAGATGAGAGAGAGAGAT 4560

Qy 4561 GCCGAGTGAATCCGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4620
 Db 4561 GCCGAGTGAATCCGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4620
 Qy 4621 GAGCAGTGAAG 4680
 Db 4621 GAGCAGTGAAG 4680
 Qy 4681 AGCTTGAAGCAGTCTGAG 4740
 Db 4681 AGCTTGAAGCAGTCTGAG 4740
 Qy 4741 ACCCATCTCACTCACTTATGAG 4800
 Db 4741 ACCCATCTCACTCACTTATGAG 4800
 Qy 4801 CTCAGCAGTGAAGTATCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
 Db 4801 CTCAGCAGTGAAGTATCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
 Qy 4861 TTCTTTGGGTCTCAGTTTCTTTCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920
 Db 4861 TTCTTTGGGTCTCAGTTTCTTTCTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920
 Qy 4921 CTAAAGTGAATGAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
 Db 4921 CTAAAGTGAATGAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
 Qy 4981 CCACAG 5040
 Db 4981 CCACAG 5040
 Qy 5041 GGCAG 5100
 Db 5041 GGCAG 5100
 Qy 5101 GGCCTCCGATCCCATCCAG 5160
 Db 5101 GGCCTCCGATCCCATCCAG 5160
 Qy 5161 TGGCTGTGTCAAG 5220
 Db 5161 TGGCTGTGTCAAG 5220
 Qy 5221 TTGGCTTTGTACAG 5280
 Db 5221 TTGGCTTTGTACAG 5280
 Qy 5281 ATTGAG 5340
 Db 5281 ATTGAG 5340
 Qy 5341 TTTACAG 5400
 Db 5341 TTTACAG 5400
 Qy 5401 ACAAAG 5460
 Db 5401 ACAAAG 5460

RESULT 2
 AF310105 5404 bp mRNA linear PRI 04-MAY-2001
 LOCUS AF310105
 DEFINITION Homo sapiens NALP1 mRNA, complete cds.
 ACCESSION AF310105
 VERSION AF310105.1 GI:11096302
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5404).
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE The pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation
JOURNAL Curr. Biol. 11 (4), R118-R120 (2001)
MEDLINE 21148093
PUBMED 11250163
REFERENCE 2 (bases 1 to 5404).
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submision
JOURNAL Submitted (28-SEP-2000) Institute of Biochemistry, University of Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
FEATURES
Source location/Qualifiers

Source
1. 5404
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
519..4808
/note="contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP1"
/protein_id="AAG30288.1"
/db_xref="GI:11096303"
translation="MAGGAKGRACVLEFLKKEIKEROLLANKAHSRSSGCTPAO
PEKTSGMVASYLVAAQVGEORAMDLATHTWOMCRLSCAQAGSAGHSPEFPSPSE
PHLSPQPTSTAVLMEPIHELPACTGSSRRVRLQPLDTSGRWREISASLYQAL
PSSPDHESPQSPNPASTAVLGWSGPPQPSLAPRQEPAGTQWPIDETSGLYTE
IRERERESKSGRPMAVAVGTVPQAFSLQPHHMPESVRESLCSPTWKNEDYFG
KFTOLLORPHRSDPLVKRSMVDYENRGLIEIRDLGEGDLDOBRITLQAG
AAGIKTLAROVKAWGRGOLYGRPOHRYVFCRSLRAGAKVSLATLIGKGTAP
APIRQILSRPRLIFILDGVDPRVNLDPSESLCHMSQOPADALISLGTITLP
EASLTITRTALQNLIPSLQARWEVLGFSSESRSKRYFRYPTDERQALRAFLVK
SNKELMALCLVPMWSMLACTLQMOQKREKLTLSKTYTLCLHYLAQALQAPLGP
QLRDLCSLAAGIWOQKTLFSPDILRKHLDGAILSTLKKGIILQEHPIPLSYFIHL
CFQEFPAASVLYLEDEKGRKHSNCTIDLETLEAYGHGFGASTFTFLGLSDEG
EREMENIFHGLSGSRNLMOWPGLSOLLQSHLSLGLCYETNKTEFLQVMAHFE
MGMCVETDMELVCTFCIKFSRHYKLOLIGRQHSRWSTPMVYLPRVVTAYWQ
ILFVLKTRNLKELDLSGNSLSHVAVSLCKTLRKPRCLLETILKACGLTADCCD
LAFGLRANQTLTELDLFSNVLTDAGAKLQRLRQPSCKLQLOVSGLSQDCLP
ASVLSASPLKELDLQNNLDVGRILCEGLRHPACILRLGDLQTLISEMDELR
ALBOEKPOLIFSRKPSVMTPEGLDGENSNSTSLRKRLGSEKASVAVANLK
LLDYKILPIELAEIABESSPEVYVPELTVPSPASQSLKRLGSDDDFMKPTPVAT
EVNDKERULRYVHPVAGSYRWPNTGLCFVNRREAVTYIEFTCDQPIEGEINPQSHM
VAGPLDIKAPGAVBAVHLPHFVALQGHVDTSLFQNAHKESEMLERKARVELHI
IVLENPSFPLGVLKIMHNAIRLIPVLSVVLVYHRAVEVTHLYLIPDCIRKE
LELCYRSGEDQLFSEFYVGHLSGIRLVQVQDKOETLMVALYKGLIMATVLLIP
ARIAPSPDLAPOLHFDYOYREOLIAVTSVEVLDLHGQVLSQEOYERVLAMTR
PSQMRKLFSLQSMRDKDGLXOALKETHHLMELMEKSKKGLPLPS"

CDS
/note="contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP1"
/protein_id="AAG30288.1"
/db_xref="GI:11096303"
translation="MAGGAKGRACVLEFLKKEIKEROLLANKAHSRSSGCTPAO
PEKTSGMVASYLVAAQVGEORAMDLATHTWOMCRLSCAQAGSAGHSPEFPSPSE
PHLSPQPTSTAVLMEPIHELPACTGSSRRVRLQPLDTSGRWREISASLYQAL
PSSPDHESPQSPNPASTAVLGWSGPPQPSLAPRQEPAGTQWPIDETSGLYTE
IRERERESKSGRPMAVAVGTVPQAFSLQPHHMPESVRESLCSPTWKNEDYFG
KFTOLLORPHRSDPLVKRSMVDYENRGLIEIRDLGEGDLDOBRITLQAG
AAGIKTLAROVKAWGRGOLYGRPOHRYVFCRSLRAGAKVSLATLIGKGTAP
APIRQILSRPRLIFILDGVDPRVNLDPSESLCHMSQOPADALISLGTITLP
EASLTITRTALQNLIPSLQARWEVLGFSSESRSKRYFRYPTDERQALRAFLVK
SNKELMALCLVPMWSMLACTLQMOQKREKLTLSKTYTLCLHYLAQALQAPLGP
QLRDLCSLAAGIWOQKTLFSPDILRKHLDGAILSTLKKGIILQEHPIPLSYFIHL
CFQEFPAASVLYLEDEKGRKHSNCTIDLETLEAYGHGFGASTFTFLGLSDEG
EREMENIFHGLSGSRNLMOWPGLSOLLQSHLSLGLCYETNKTEFLQVMAHFE
MGMCVETDMELVCTFCIKFSRHYKLOLIGRQHSRWSTPMVYLPRVVTAYWQ
ILFVLKTRNLKELDLSGNSLSHVAVSLCKTLRKPRCLLETILKACGLTADCCD
LAFGLRANQTLTELDLFSNVLTDAGAKLQRLRQPSCKLQLOVSGLSQDCLP
ASVLSASPLKELDLQNNLDVGRILCEGLRHPACILRLGDLQTLISEMDELR
ALBOEKPOLIFSRKPSVMTPEGLDGENSNSTSLRKRLGSEKASVAVANLK
LLDYKILPIELAEIABESSPEVYVPELTVPSPASQSLKRLGSDDDFMKPTPVAT
EVNDKERULRYVHPVAGSYRWPNTGLCFVNRREAVTYIEFTCDQPIEGEINPQSHM
VAGPLDIKAPGAVBAVHLPHFVALQGHVDTSLFQNAHKESEMLERKARVELHI
IVLENPSFPLGVLKIMHNAIRLIPVLSVVLVYHRAVEVTHLYLIPDCIRKE
LELCYRSGEDQLFSEFYVGHLSGIRLVQVQDKOETLMVALYKGLIMATVLLIP
ARIAPSPDLAPOLHFDYOYREOLIAVTSVEVLDLHGQVLSQEOYERVLAMTR
PSQMRKLFSLQSMRDKDGLXOALKETHHLMELMEKSKKGLPLPS"

BASE COUNT 1284 a 1518 c 1467 g 1135 t
ORIGIN
Query Match 99.2%; Score 5400.8; DB 9; Length 5404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5402; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 CAGGGCTGGAAGAGTCTGGAAGAACTGGAGACCAAGCCCGGGGCTTCACTTGGGT 64
DB 1 CAGGGCTGGAAGAGTCTGGAAGAACTGGAGACCAAGCCCGGGGCTTCACTTGGGT 60
QY 65 TCTGAAGCCCATTCCTGCTCTGCGGCTCTCCCAACCCCACTTCTCAAGCCTTGGAG 124
DB 61 TCTGAAGCCCATTCCTGCTCTGCGGCTCTCCCAACCCCACTTCTCAAGCCTTGGAG 120
QY 125 CTCAAGGCTTATCTCAGAGTCCAGAGCCCAAGAGAGGAGAAATCTGAGAAACACAGA 184
DB 121 CTCAAGGCTTATCTCAGAGTCCAGAGCCCAAGAGAGGAGAAATCTGAGAAACACAGA 180
QY 185 ACAAGTGAAGCTTGGCCCAACCCCATCTCCGTCACACATCTCCCTTCACTTCACTTC 244
DB 181 ACAAGTGAAGCTTGGCCCAACCCCATCTCCGTCACACATCTCCCTTCACTTCACTTC 240
QY 245 CCTGAGCTGAGCCCATCCAGAGACCTCCCTATCAGTGAATCTTCAAGTGTCT 304

DB 241 CCTGAGCTGAGCCCATCCAGAGACCTCCCTATCAGTGAATCTTCAAGTGTCT 300
QY 305 TGAAGGCCCCCTGAGGCTCTCCCTCCCTGAGCTTTTCTCAACAATCCCTATCGGC 364
DB 301 TGAAGGCCCCCTGAGGCTCTCCCTCCCTGAGCTTTTCTCAACAATCCCTATCGGC 360
QY 365 GTCTATCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
DB 361 GTCTATCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 425 GTAAAGCCAGAGGAGAAAGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
DB 421 GTAAAGCCAGAGGAGAAAGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 485 GAACATCCCCCAGACACCTCTTAACTCCGAGACAGAGATGCTGCTGCTGCTGCTGCTGCT 544
DB 481 GAACATCCCCCAGACACCTCTTAACTCCGAGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 545 GCTGAGCTGCTTATCTGAGATCTTGAAGAGAGAGCTGAGAGATTCAGCTTCTGC 604
DB 541 GCTGAGCTGCTTATCTGAGATCTTGAAGAGAGAGCTGAGAGATTCAGCTTCTGC 600
QY 605 TGGCAATTAAGGAGCACTCCAGAGCTTTCGGTGAAGACCCGCTCAGCCAGAGAGA 664
DB 601 TGGCAATTAAGGAGCACTCCAGAGCTTTCGGTGAAGACCCGCTCAGCCAGAGAGA 660
QY 665 CGAGTGGCATGAGATGAGCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
DB 661 CGAGTGGCATGAGATGAGCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 725 ACCTAGCCTCTCATACCTGAGAGAGATGAGGCTGAGGTCACTGTGCGCCCAAGCCACG 784
DB 721 ACCTAGCCTCTCATACCTGAGAGAGATGAGGCTGAGGTCACTGTGCGCCCAAGCCACG 780
QY 785 AAGGGGAGGAGCACTCCCTCCCTCATTCGCCCTACAGCCCAAGGAAACCCACCTGAGGCTTC 844
DB 781 AAGGGGAGGAGCACTCCCTCCCTCATTCGCCCTACAGCCCAAGGAAACCCACCTGAGGCTTC 840
QY 845 CCAAGCCCAACCACTCCACCGCAGTGTAAATCCCTGATTCATGAATTCGCGCGGGGT 904
DB 841 CCAAGCCCAACCACTCCACCGCAGTGTAAATCCCTGATTCATGAATTCGCGCGGGGT 900
QY 905 GCATCCAGGGCTCAGAGAGAGGGTTTGAAGACAGCTGCTGACACATCTGAGAGCGGCT 964
DB 901 GCATCCAGGGCTCAGAGAGAGGGTTTGAAGACAGCTGCTGACACATCTGAGAGCGGCT 960
QY 965 GGAAGAAATCTGAGCTCCTCCTCTTACCAAGCTCTTCCAAAGCTCCCAAGCATGAGT 1024
DB 961 GGAAGAAATCTGAGCTCCTCCTCTTACCAAGCTCTTCCAAAGCTCCCAAGCATGAGT 1020
QY 1025 CTCCAGCCAGAGATCAACCAACGCCCCCATCATCCAGCAGTGTGAGGAGCTGAGGAT 1084
DB 1021 CTCCAGCCAGAGATCAACCAACGCCCCCATCATCCAGCAGTGTGAGGAGCTGAGGAT 1080
QY 1085 CCCCACCTCAGCCAGCTTACACCCAGAGAGAGAGGCTCTGAGAGCCCAATGAGCTTC 1144
DB 1081 CCCCACCTCAGCCAGCTTACACCCAGAGAGAGAGGCTCTGAGAGCCCAATGAGCTTC 1140
QY 1145 TGAATGAAGATGAGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAAATCAG 1204
DB 1141 TGAATGAAGATGAGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAAATCAG 1200
QY 1205 AGAAAGCAGGCCCCCAGTGGGACAGCGGTGAGAAACGCCCCCAGAGCGCAGACAGCC 1264
DB 1201 AGAAAGCAGGCCCCCAGTGGGACAGCGGTGAGAAACGCCCCCAGAGCGCAGACAGCC 1260
QY 1265 TACAGCCCAACCAACCAATGAGAGCTTCTGAGAGAGAGAGCTGCTGCTTCCCATGAGC 1324
DB 1261 TACAGCCCAACCAACCAATGAGAGCTTCTGAGAGAGAGAGCTGCTGCTTCCCATGAGC 1320
QY 1325 CTTGAGAAATGAGAGATTTTAAACCAAAATTCACACAGCTGCTACTTCAAAAGACCTC 1384
DB 1321 CTTGAGAAATGAGAGATTTTAAACCAAAATTCACACAGCTGCTACTTCAAAAGACCTC 1380

QY	1385	ACCCAGAAAGCCAGATCCCTGGTCAAGAAAGCTGGCTGATTTATGTGAGAGAAATC	1444A
Db	1381	ACCCAGAAAGCCAGATCCCTGGTCAAGAAAGCTGGCTGATTTATGTGAGAGAAATC	1440A
QY	1445	GAGACATTTTAAATGAGATCAGAGACTAATTTGGCCAGGCTGGATACCAAGAACTC	1504A
Db	1441	GAGACATTTTAAATGAGATCAGAGACTAATTTGGCCAGGCTGGATACCAAGAACTC	1500A
QY	1505	GCATATGTCATCTGCAAGGGGGCTGTGGAAATTGGGAATCAACATGGCCACAGCAAGTGA	1564A
Db	1501	GCATATGTCATCTGCAAGGGGGCTGTGGAAATTGGGAATCAACATGGCCACAGCAAGTGA	1560A
QY	1565	AGGAAGCCTGGGGGGAGAGCCAGCTGTATGGGGAAACGGTTCCAGATGTCTTATCTTCA	1624A
Db	1561	AGGAAGCCTGGGGGGAGAGCCAGCTGTATGGGGAAACGGTTCCAGATGTCTTATCTTCA	1620A
QY	1625	GCTGCAGAGAGCTGGCCCAAGTCCAAGTGTAGATGTCTGCTGAGCTCATCGGAAAAAGATG	1684A
Db	1621	GCTGCAGAGAGCTGGCCCAAGTCCAAGTGTAGATGTCTGCTGAGCTCATCGGAAAAAGATG	1680A
QY	1685	GGAACAGCATCTCGGGCTCCCATTTAGACAGATCTGTCTTAGGCCAGAGGGGCTGCTTTCA	1744A
Db	1681	GGAACAGCATCTCGGGCTCCCATTTAGACAGATCTGTCTTAGGCCAGAGGGGCTGCTTTCA	1740A
QY	1745	TCCTCGATGGGTATGATGAGCCAGGATGGGTCTTGACAGAGCCGAGTCTGAGCTCATC	1804A
Db	1741	TCCTCGATGGGTATGATGAGCCAGGATGGGTCTTGACAGAGCCGAGTCTGAGCTCATC	1800A
QY	1805	TGCACTGAGAGCCAGCCACAGCCGGCGAGTCACTGTGGGACGATTTGCTGGGGAAAACTA	1864A
Db	1801	TGCACTGAGAGCCAGCCACAGCCGGCGAGTCACTGTGGGACGATTTGCTGGGGAAAACTA	1860A
QY	1865	TACTTCCCGAAGGACCTCTCTGATCAAGGCTCGGACCAACAGCTCTGAGAAACCTCATTC	1924A
Db	1861	TACTTCCCGAAGGACCTCTCTGATCAAGGCTCGGACCAACAGCTCTGAGAAACCTCATTC	1920A
QY	1925	CTTCTTTGGAGCAGGACGTTGGGTAGAGGTCTTGAGGTTCTCTGAGTCCAGCAGGAAAG	1984A
Db	1921	CTTCTTTGGAGCAGGACGTTGGGTAGAGGTCTTGAGGTTCTCTGAGTCCAGCAGGAAAG	1980A
QY	1985	AATATTTCTACAGATATTTCAAGATGAAAGCAAGCAATTAGACCCTTTAGGTTGGTCA	2044A
Db	1981	AATATTTCTACAGATATTTCAAGATGAAAGCAAGCAATTAGACCCTTTAGGTTGGTCA	2040A
QY	2045	AATCAAAACAAAGAGCTGGGGCCCTGTATCTTGAGCCCTGGGTGTCCTGGCTGGCTGCA	2104A
Db	2041	AATCAAAACAAAGAGCTGGGGCCCTGTATCTTGAGCCCTGGGTGTCCTGGCTGGCTGCA	2100A
QY	2105	CTTGCTGTAGTCAGCAGATGAAAGCGGAAGAAAACTCACTGATCTTCCAAGACCAACA	2164A
Db	2101	CTTGCTGTAGTCAGCAGATGAAAGCGGAAGAAAACTCACTGATCTTCCAAGACCAACA	2160A
QY	2165	CAACCCCTGTGTCACTATTAACCTTGCCCAAGGCTGCCCAAGCTCAGGCATTTGGGACCCAGC	2224A
Db	2161	CAACCCCTGTGTCACTATTAACCTTGCCCAAGGCTGCCCAAGCTCAGGCATTTGGGACCCAGC	2220A
QY	2225	TCAGAGACCTCTGTCTCTGTGCTGTGAGGGGCATCTGGCAAAAAAGACCCCTTTTCAAGTC	2284A
Db	2221	TCAGAGACCTCTGTCTCTGTGCTGTGAGGGGCATCTGGCAAAAAAGACCCCTTTTCAAGTC	2280A
QY	2285	CAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATATCTCCAACCTTCTTGAAGATGG	2344A
Db	2281	CAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATATCTCCAACCTTCTTGAAGATGG	2340A
QY	2345	GTAATCTTCAAGAGACCCCAATCCCTCGAGCTACAGCTTCAATCAACCTGTGTTTCCAG	2404A
Db	2341	GTAATCTTCAAGAGACCCCAATCCCTCGAGCTACAGCTTCAATCAACCTGTGTTTCCAG	2400A
QY	2405	AGTCTTTTGACAGCATGTCTATGTCTTGAAGATGAGAGGGGAGAGTAAACATTTCTA	2464A
Db	2401	AGTCTTTTGACAGCATGTCTATGTCTTGAAGATGAGAGGGGAGAGTAAACATTTCTA	2460A

QY	2465	TTTGCACTCATGATTTTGGAAAAAGCGGTAGAAAGATATGGAATTCATGGCCCTGTTTGGG	2524
Db	2461	ATTGCATCATAGATTTTGGAAAAAGCGTAGAAGCATATGGAATTCATGGCCCTGTTTGGG	2520
QY	2525	-CATGACCAACACCGTTTCTATTGGGCGCTGTTAAAGTAGTAGAGGGGAGAGAGATGAGAG	2584
Db	2521	CATCAACCAACCGTTTCTATTGGGCGCTGTTAAAGTAGTAGAGGGGAGAGAGATGAGAG	2580
QY	2585	ACATCTTTCACTGCCGCGTGTCTTAGGGAGAGAACTGATATCAGTGGGTCCCGTCCCTGC	2644
Db	2581	ACATCTTTCACTGCCGCGTGTCTTAGGGAGAGAACTGATATCAGTGGGTCCCGTCCCTGC	2640
QY	2645	AGCTGCTGCTGACAGCAACTCTGTGAGAGTCCCTCACTGCTGTAGAGAGCTGGAAACA	2704
Db	2641	AGCTGCTGCTGACAGCAACTCTGTGAGAGTCCCTCACTGCTGTAGAGAGCTGGAAACA	2700
QY	2705	AAACGTTTCTGACACAAATGATGATGCCATTTGGAAGAAATGGGCGATGTGTAGAAAACAG	2764
Db	2701	AAACGTTTCTGACACAAATGATGATGCCATTTGGAAGAAATGGGCGATGTGTAGAAAACAG	2760
QY	2765	ACATGGAGCTTTATGATGACATTTTCGCAATTAATCAGCCGCGACGTGAGAAAGCTTC	2824
Db	2761	ACATGGAGCTTTATGATGACATTTTCGCAATTAATCAGCCGCGACGTGAGAAAGCTTC	2820
QY	2825	AGCTGATTGAGGGGACAGGACAGACAGATCAACATGAGACCCCAACCATGATGTCTGTGTCA	2884
Db	2821	AGCTGATTGAGGGGACAGGACAGACAGATCAACATGAGACCCCAACCATGATGTCTGTGTCA	2880
QY	2885	GGTGGGTCCCACTGACAGATATGCTTATTTGGCAGATTTCTTTCTCCGTCTCAAGTCAACA	2944
Db	2881	GGTGGGTCCCACTGACAGATATGCTTATTTGGCAGATTTCTTTCTCCGTCTCAAGTCAACA	2940
QY	2945	GAACCTGGAAGAGACTGACCTTAAGTGTGAAACTGCGTGAAGCACTCTGCAGTGAAGAGTC	3004
Db	2941	GAACCTGGAAGAGACTGACCTTAAGTGTGAAACTGCGTGAAGCACTCTGCAGTGAAGAGTC	3000
QY	3005	TTTGTAAAGCCCTGAGACGCGCTTCGCTGCTCTTGAAGACCTTCGCGTGGCTGGCTGTG	3064
Db	3001	TTTGTAAAGCCCTGAGACGCGCTTCGCTGCTCTTGAAGACCTTCGCGTGGCTGGCTGTG	3060
QY	3065	GCCCTCAGACGTGAAGACATGCAAGGACCTTCGCTTTGGGCTGAGAGGCCAACAGACCTTGA	3124
Db	3061	GCCCTCAGACGTGAAGACATGCAAGGACCTTCGCTTTGGGCTGAGAGGCCAACAGACCTTGA	3120
QY	3125	CCGAGCTG3ACTTAGGCTTCAATATGTGCTCAACGATGTGTGAGGCCAAACACTTGTGCCAGA	3184
Db	3121	CCGAGCTG3ACTTAGGCTTCAATATGTGCTCAACGATGTGTGAGGCCAAACACTTGTGCCAGA	3180
QY	3185	GACTGAGACAGCCGAGCTGCAAGTTCAGAGGACCTGCACTGTGTATGCTGTGGCTTCAAGT	3244
Db	3181	GACTGAGACAGCCGAGCTGCAAGTTCAGAGGACCTGCACTGTGTATGCTGTGGCTTCAAGT	3240
QY	3245	CTGACTGTGTCAGAGACTGTGGCTCTCTGTGCTTAATGTGCAGCGCCAGCGCTGAAGAGCTAG	3304
Db	3241	CTGACTGTGTCAGAGACTGTGGCTCTCTGTGCTTAATGTGCAGCGCCAGCGCTGAAGAGCTAG	3300
QY	3305	ACCTGACAGCAAAACAACCTGTGATGACGTTGGCGTGTGCGACCTGTCTGTAGGGGCTCAAGCC	3364
Db	3301	ACCTGACAGCAAAACAACCTGTGATGACGTTGGCGTGTGCGACCTGTCTGTAGGGGCTCAAGCC	3360
QY	3365	ATCCTGCTGTGAACCTATACGCTTGCGGGCTGTGACCAAGACAACCTATAGATATGAGATGA	3424
Db	3361	ATCCTGCTGTGAACCTATACGCTTGCGGGCTGTGACCAAGACAACCTATAGATATGAGATGA	3420
QY	3425	GGCAGGAACCTGAGGGGCGCTGTGAGCAGAGAAACCTCAGCTGTCTCATCTTCAAGCAACGGA	3484
Db	3421	GGCAGGAACCTGAGGGGCGCTGTGAGCAGAGAAACCTCAGCTGTCTCATCTTCAAGCAACGGA	3480
QY	3485	AAACCAATGTATATACCCCTCATCTGAGGGGCTGTGATACGGAGAGATATGATATATGACAT	3544
Db	3481	AAACCAATGTATATACCCCTCATCTGAGGGGCTGTGATACGGAGAGATATGATATATGACAT	3540
QY	3545	CCTCACTCAAGCGGCAAGACTCGGATTCAGAGAGGGGCGGCTTCCCATGTTGCTCAGGCTTA	3604

Db 3541 CCTCACTCAAGCGGACAGAGACTCGGATCAAGAGGGGGCTTCCCATGTTGCTCAGGCTA 3600
 QY 3605 ATCTCAAACTCTGAGAGTGAACAAGATCTTCCCAATGCTGAGATTGCAAGGAAAGCT 3664
 Db 3601 ATCTCAAACTCTGAGAGTGAACAAGATCTTCCCAATGCTGAGATTGCAAGGAAAGCT 3660
 QY 3665 CCCCCAGAGGATGATCCGGTGAATCTTGTGCGTCTTCTCTGCTCTTCAAGGGGACC 3724
 Db 3661 CCCCAGAGGATGATCCGGTGAATCTTGTGCGTCTTCTCTGCTCTTCAAGGGGACC 3720
 QY 3725 TGCAATAGAAAGCTTTGGGGCTGACGATGACTTGGGGGGCCCAAGGGGGCTGAGCTA 3784
 Db 3721 TGCAATAGAAAGCTTTGGGGCTGACGATGACTTGGGGGGCCCAAGGGGGCTGAGCTA 3780
 QY 3785 CTGAGGATGATGACAAAGAAAGAACTTGTACGAGTTCACTTCCCTGAGCTGGCTCT 3844
 Db 3781 CTGAGGATGATGACAAAGAAAGAACTTGTACGAGTTCACTTCCCTGAGCTGGCTCT 3840
 QY 3845 ACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGAGAAAGCGGTGACCGTTGAGATTG 3904
 Db 3841 ACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGAGAAAGCGGTGACCGTTGAGATTG 3900
 QY 3905 AATTCTGTGATGGGACCAAGTCTGGGTGATGATCAACCAAGCAGAGCTGATGGATTG 3964
 Db 3901 AATTCTGTGATGGGACCAAGTCTGGGTGATGATCAACCAAGCAGAGCTGATGGATTG 3960
 QY 3965 CAGGGCTCTGCTGAGATCAAGGCTGAGCTGAGCTGAGAAAGCTGTGACCTCCCTC 4024
 Db 3961 CAGGGCTCTGCTGAGATCAAGGCTGAGCTGAGCTGAGAAAGCTGTGACCTCCCTC 4020
 QY 4025 ACTTGTGCTCTCTCAAGGGGGCCATGTGACACATCCCTGTTCCAAATGGGCCACTTTA 4084
 Db 4021 ACTTGTGCTCTCTCAAGGGGGCCATGTGACACATCCCTGTTCCAAATGGGCCACTTTA 4080
 QY 4085 AAGAGGAGGGGATGCTCTGGAAGAGCAGCAGGGGTGAGCTGCATCACTAGTTGTG 4144
 Db 4081 AAGAGGAGGGGATGCTCTGGAAGAGCAGCAGGGGTGAGCTGCATCACTAGTTGTG 4140
 QY 4145 AAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAAATGATCATATGCTCTGCGCT 4204
 Db 4141 AAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAAATGATCATATGCTCTGCGCT 4200
 QY 4205 TCATTCCTGCTGCTCTGTGTGTGCTTTACACCGCGTCCATCTGAGAGAGTCACT 4264
 Db 4201 TCATTCCTGCTGCTCTGTGTGTGCTTTACACCGCGTCCATCTGAGAGAGTCACT 4260
 QY 4265 TCACCTCTACCTGATCCCAAGTGACTGCTCCATTCGGAAGGAACTGAGCTGCTATC 4324
 Db 4261 TCACCTCTACCTGATCCCAAGTGACTGCTCCATTCGGAAGGAACTGAGCTGCTATC 4320
 QY 4325 GAAGCCCTGGAAGAGCAAGCTGTTCTCGAGTTCTACGTTGSCCACTTGGGATCAGGGA 4384
 Db 4321 GAAGCCCTGGAAGAGCAAGCTGTTCTCGAGTTCTACGTTGSCCACTTGGGATCAGGGA 4380
 QY 4385 TCAGGCTGCAAGTGAAGACAAAGAAAGATGAGACTTGTGTGTGAGAGCCTTGTGAAAC 4444
 Db 4381 TCAGGCTGCAAGTGAAGACAAAGAAAGATGAGACTTGTGTGTGAGAGCCTTGTGAAAC 4440
 QY 4445 CAGAGATCTCAGGCTGCAACTCTGATCCCTCAGCCCGGATGAGCGTACCTTCAAC 4504
 Db 4441 CAGAGATCTCAGGCTGCAACTCTGATCCCTCAGCCCGGATGAGCGTACCTTCAAC 4500
 QY 4505 CTCTGATGCCCCGAGCTTGTGCTGCACTTGTGACAGATGATGAGAGCAGCTGATAGCCC 4564
 Db 4501 CTCTGATGCCCCGAGCTTGTGCTGCACTTGTGACAGATGATGAGAGCAGCTGATAGCCC 4560
 QY 4565 GAGTGAATCTGCTGAGAGTGTCTTGTGACAAACTGATGAGCAGAGTGTGAGCAAGAAC 4624
 Db 4561 GAGTGAATCTGCTGAGAGTGTCTTGTGACAAACTGATGAGCAGAGTGTGAGCAAGAAC 4620
 QY 4625 AGTCAAGAGGGTCTGGCTGAGAAACAGAGGCCAGCAATGCGGAGGATGTTCAAGCT 4684

Db 4621 AGTCAAGAGGGTCTGGCTGAGAAACAGAGGCCCAAGCCAGATGCGGAAGCTGTTCAAGCT 4680
 QY 4685 TGAAGCACTCTGGAGCCGGAGTGCAGAAAGATGAGACTTACCAAGCTCTGAAGAGAACCC 4744
 Db 4681 TGAAGCACTCTGGAGCCGGAGTGCAGAAAGATGAGACTTACCAAGCCCTGAAGAGAACCC 4740
 QY 4745 ATCTCACTCATTTAATGGAATCTGTGGAGAAAGGCAAGGAAAGGAACTCTGCACTCA 4804
 Db 4741 ATCTCACTCATTTAATGGAATCTGTGGAGAAAGGCAAGGAAAGGAACTCTGCACTCA 4800
 QY 4805 GCAGCTGAAGTATCAACACAGCCCTTGAACCTTTGAGTCTGCTGGCTGAGACCTTCT 4864
 Db 4801 GCAGCTGAAGTATCAACACAGCCCTTGAACCTTTGAGTCTGCTGGCTGAGACCTTCT 4860
 QY 4865 TTGGGTCTCAGTTTCTTCTCTGCAAAAGATGTCATCTGTTGCTTCCAGACTTAA 4924
 Db 4861 TTGGGTCTCAGTTTCTTCTCTGCAAAAGATGTCATCTGTTGCTTCCAGACTTAA 4920
 QY 4925 AGTAAATGGAATTTGATGATGCTTTGTGGGCAATTAATGTCTCATATGCCAGGATGCGAC 4984
 Db 4921 AGTAAATGGAATTTGATGATGCTTTGTGGGCAATTAATGTCTCATATGCCAGGATGCGAC 4980
 QY 4985 AGGGGGCCCCAGTCCAGGTGGCTTAACAGAGCATCTCAGGGAAATGTCATCTGAGCTGGCA 5044
 Db 4981 AGGGGGCCCCAGTCCAGGTGGCTTAACAGAGCATCTCAGGGAAATGTCATCTGAGCTGGCA 5040
 QY 5045 AGACCCCTGCAAGCTTCATATGAGCTTCATCTGTGTGGCCACAGCAGCCCAAGCTTGAAGCCC 5104
 Db 5041 AGACCCCTGCAAGCTTCATATGAGCTTCATCTGTGTGGCCACAGCAGCCCAAGCTTGAAGCCC 5100
 QY 5105 TCCGATATCCATCATGAGCCGCAAGAGAAATAGAGAGGAACTATGAAACCATTTGCTCGGC 5164
 Db 5101 TCCGATATCCATCATGAGCCGCAAGAGAAATAGAGAGGAACTATGAAACCATTTGCTCGGC 5160
 QY 5165 TGTGTCAAGGATGAGCCCAAAATTTGGGTTGAGCTGAGGAGGCCAGTGGATTTCTTGG 5224
 Db 5161 TGTGTCAAGGATGAGCCCAAAATTTGGGTTGAGCTGAGGAGGCCAGTGGATTTCTTGG 5220
 QY 5225 CTTTGTACAGAAAGATCTACAGAGCAAGCCAAAGATGAAGTGAAGAGATTATTC 5284
 Db 5221 CTTTGTACAGAAAGATCTACAGAGCAAGCCAAAGATGAAGTGAAGAGATTATTC 5280
 QY 5285 AGAAAAATTAAGAGATCAAGCTCTTTTGAATTTGTCTAGAGAGGTTTCCAGTTTAA 5344
 Db 5281 AGAAAAATTAAGAGATCAAGCTCTTTTGAATTTGTCTAGAGAGGTTTCCAGTTTAA 5340
 QY 5345 CCAAGAAAACCCCTATTAATTAATAATTTTAACTTAATTAAGATTAATAATAATCAAA 5404
 Db 5341 CCAAGAAAACCCCTATTAATTAATAATTTTAACTTAATTAAGATTAATAATAATAATAA 5400
 QY 5405 AAAA 5408
 Db 5401 AAAA 5404

RESULT 3
 AX459865
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Techopp, J. and Martinon, F.
 Proteins and dna sequences underlying these proteins used for
 treating inflammations
 Patent: WO 0240668-A 8 23-MAY-2002;
 Apotech Research and Development Ltd. (CH)
 Location/Qualifiers

```

source                1. 5100
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /note="NALP1.cdna"
BASE COUNT            1251 a 1374 c 1406 g 1069 t
ORIGIN
Query Match          89.7%; Score 4882.4; DB 6; Length 5100;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 516 GACGAGATGGCTGCGGAGCGCTGGGGCCGCTGGCTGTTACTTGTGAAGTTCCGGAAGA 575
DB 153 GAACAAATGGCTGGCGAGCTGGGGCCGCTGGCTGTTACTTGTGAAGTTCCGGAAGA 212
QY 576 GAGAGAGCTGAAGAGTTCCAGCTTGTGCTGCGCAATTAAGCGACTCCAGAGACTCTTC 635
DB 213 GAGAGAGCTGAAGAGTTCCAGCTTGTGCTGCGCAATTAAGCGACTCCAGAGACTCTTC 272
QY 636 GGGTGAACACCCGCTCAGCCAGAGAAAGAGATGGCATGGAGTGGCTCTGTAAGTGT 695
DB 273 GGGTGAACACCCGCTCAGCCAGAGAAAGAGATGGCATGGAGTGGCTCTGTAAGTGT 332
QY 696 GGGCTAGTAAGGGAGAGAGCGGGGCTGGGACTAGCCCTCCATACCTGGAGAGAGATGGG 755
DB 333 GGGCTAGTAAGGGAGAGAGCGGGGCTGGGACTAGCCCTCCATACCTGGAGAGAGATGGG 392
QY 756 GCTGAGTCACTGTGCGCCCAAGGCCAGAGAGGGGAGGCGCACTCTCCCTCATTCCTTA 815
DB 393 GCTGAGTCACTGTGCGCCCAAGGCCAGAGAGGGGAGGCGCACTCTCCCTCATTCCTTA 452
QY 816 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTAAAT 875
DB 453 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTAAAT 512
QY 876 GGGCTGATCCATTAATTTGCCGGGGGGGTGCAACCAAGGGCTCAAGAGAGAGGTTTGG 935
DB 513 GGGCTGATCCATTAATTTGCCGGGGGGGTGCAACCAAGGGCTCAAGAGAGAGGTTTGG 572
QY 936 ACAAGCTCCTGACACATCTGAGACCGCGCTGAGAGAAATCTCTGCTCACTCTCTACA 995
DB 573 ACAAGCTCCTGACACATCTGAGACCGCGCTGAGAGAAATCTCTGCTCACTCTCTACA 632
QY 996 AGCTCTTCCAAAGCTCCCAAGACATGAGTCTCAAGCCAGAGATCAACCAAGCCCCAC 1055
DB 633 AGCTCTTCCAAAGCTCCCAAGACATGAGTCTCAAGCCAGAGATCAACCAAGCCCCAC 692
QY 1056 ATCCACAGCAGTGTGGGGAGCTGGGATCCCAACCTCAAGCCCAAGCTTACAGCAAG 1115
DB 693 ATCCACAGCAGTGTGGGGAGCTGGGATCCCAACCTCAAGCCCAAGCTTACAGCAAG 752
QY 1116 GCAAGAGGCTCTGAGGACCCCAATGGCTCTGATGAAACGTCAAGAAATTTTACTACAGA 1175
DB 753 GCAAGAGGCTCTGAGGACCCCAATGGCTCTGATGAAACGTCAAGAAATTTTACTACAGA 812
QY 1176 AATGAGAGAAAGAGAGAGAAATCAAGAAAGCGGCCCTCCATGGGCAAGCGGTGGT 1235
DB 813 AATGAGAGAAAGAGAGAGAAATCAAGAAAGCGGCCCTCCATGGGCAAGCGGTGGT 872
QY 1236 AGGAAGCCCCCAAGGCGGACACAGCTACAGCCCCACACACCAACCATGGAGGCTTC 1295
DB 873 AGGAAGCCCCCAAGGCGGACACAGCTACAGCCCCACACACCAACCATGGAGGCTTC 932
QY 1296 TGTGAGAGAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACAAAATT 1355
DB 933 TGTGAGAGAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACAAAATT 992
QY 1356 CACGAGCTGCTACTTCTTACAAAGACTCAACCCCAAGAGCCAAAGTCCCTGCTCAAG 1415
DB 993 CACGAGCTGCTACTTCTTACAAAGACTCAACCCCAAGAGCCAAAGTCCCTGCTCAAG 1052
QY 1416 AAGCTGAGCTGATATGTGAGAGAGAAATGAGGACATTTAATTGAGATCAAGACTTATT 1475

```

```

DB 1053 AAGCTGAGCTGATTAATGTGAGAGAAATCGAGACATTTAATTGAGATCAAGACTTATT 1112
QY 1476 TGGCCCAAGGCTGAGATACCCAAAGACTCCGATAGTCAATCTGAGAGGGGCTGCTGAAT 1535
DB 1113 TGGCCCAAGGCTGAGATACCCAAAGACTCCGATAGTCAATCTGAGAGGGGCTGCTGAAT 1172
QY 1536 TGGGAAGTCAACACTGGCCAGGCAAGTGAAGAAAGCTGGGGAGAGGCCAGCTGTATGG 1595
DB 1173 TGGGAAGTCAACACTGGCCAGGCAAGTGAAGAAAGCTGGGGAGAGGCCAGCTGTATGG 1232
QY 1596 GGAACCGCTTCACAGATGCTTCTTACTTCACTGACAGAGAGCTGGCCAGTCCAAGGTGT 1655
DB 1233 GGAACCGCTTCACAGATGCTTCTTACTTCACTGACAGAGAGCTGGCCAGTCCAAGGTGT 1292
QY 1656 GAGTCTGCTGAGCTCATCCGAAAAAGATGGAGAGCCACTCCGCTCCATTAGACAGAT 1715
DB 1293 GAGTCTGCTGAGCTCATCCGAAAAAGATGGAGAGCCACTCCGCTCCATTAGACAGAT 1352
QY 1716 CTTGTCTAGGCCAGAGGGGCTGCTCTTCACTCTGATGATGATGATGATGATGATGATG 1775
DB 1353 CTTGTCTAGGCCAGAGGGGCTGCTCTTCACTCTGATGATGATGATGATGATGATGATG 1412
QY 1776 CTTGAGAGAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCAGCAAGCCGCGGATGC 1835
DB 1413 CTTGAGAGAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCAGCAAGCCGCGGATGC 1472
QY 1836 ACTGCTGGGAGTTTGTGGGAAAACTATACTTCCGAGGCACTCTCTGATCAAGGC 1895
DB 1473 ACTGCTGGGAGTTTGTGGGAAAACTATACTTCCGAGGCACTCTCTGATCAAGGC 1532
QY 1896 TCGGACCAACGCTCTGCAAAACCTCATTTCTTTTGAAGCAGCAGCTGGGTGAAGGT 1955
DB 1533 TCGGACCAACGCTCTGCAAAACCTCATTTCTTTTGAAGCAGCAGCTGGGTGAAGGT 1592
QY 1956 CTTGGGGTCTCTGAGTCCAGAGCAAGAAAGAAATTTTCTACAGATATTCTCAAGATGAAG 2015
DB 1593 CTTGGGGTCTCTGAGTCCAGAGCAAGAAAGAAATTTTCTACAGATATTCTCAAGATGAAG 1652
QY 2016 GCAAGCAATTAGAAGCTTTAGTTGTGTAATCAAAACAAAGAGCTTGAGGCTGTGTCT 2075
DB 1653 GCAAGCAATTAGAAGCTTTAGTTGTGTAATCAAAACAAAGAGCTTGAGGCTGTGTCT 1712
QY 2076 TGTGCTCTGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2135
DB 1713 TGTGCTCTGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1772
QY 2136 AAAAAGTCACTGACTTCCAAAGCAACCAACCAACCTCTGCTGCTGCTGCTGCTGCTGCTG 2195
DB 1773 AAAAAGTCACTGACTTCCAAAGCAACCAACCAACCTCTGCTGCTGCTGCTGCTGCTGCTG 1832
QY 2196 TCTTCAAGCTCAGCCATTTGGGACCCCAAGCTCAGAGACTCTGCTGCTGCTGCTGAGG 2255
DB 1833 TCTTCAAGCTCAGCCATTTGGGACCCCAAGCTCAGAGACTCTGCTGCTGCTGCTGAGG 1892
QY 2256 CATCTGGCAAAAAAGACCTTTTCACTGCAATGACCTTCAAGAGAGATGGTTAAGTGG 2315
DB 1893 CATCTGGCAAAAAAGACCTTTTCACTGCAATGACCTTCAAGAGAGATGGTTAAGTGG 1952
QY 2316 GGCATCATCTCCAGCTTCTTGAAGATGGTATTTCTTCAAGAGCAACCCATCCCTGAG 2375
DB 1953 GGCATCATCTCCAGCTTCTTGAAGATGGTATTTCTTCAAGAGCAACCCATCCCTGAG 2012
QY 2376 CTAAGCTTATTCACCTCTGTTTCCAAAGTTCTTTGACGAATGCTATGTCTTGA 2435
DB 2013 CTAAGCTTATTCACCTCTGTTTCCAGAGTTCTTTGACGAATGCTATGTCTTGA 2072
QY 2436 GGAATGAAGAGGAGAGTAAACATTTAATGATATGATTTGGAAGAAAGAGCTAGA 2495
DB 2073 GGAATGAAGAGGAGAGTAAACATTTAATGATATGATTTGGAAGAAAGAGCTAGA 2132
QY 2496 AGCATATGAATATCAGGCTGTTTGGGAGATCAACCAAGCTTTCTATTTGGGCTGTT 2555

```

Db 2133 AGCATATGGAATACATGCGCTGTTTGGGGCATCAACACAGTTTCTATTTGGGCTGT 2192
Qy 2556 AAGGATGAGGGGGAGAGAGAGATGAGAAACATCTTTCACTGCCGGCTGTCTCAGGGAG 2615
Db 2193 AAGGATGAGGGGGAGAGAGAGATGAGAAACATCTTTCACTGCCGGCTGTCTCAGGGAG 2252
Qy 2616 GAACCTGATGAGTGGGTCCCGCTCCTGAGCTGTGTGAGGACACACTCTGAGAGTC 2675
Db 2253 GAACCTGATGAGTGGGTCCCGCTCCTGAGCTGTGTGAGGACACACTCTGAGAGTC 2312
Qy 2676 CCTCCACTGCTTGTATCGAGACTCGGAAACAAACGTTCTGTGACACAGTGTAGGCCATT 2735
Db 2213 CCTCCACTGCTTGTATGAGACTCGGAAACAAACGTTCTGTGACACAGTGTAGGCCATT 2372
Qy 2736 CGAAGAAATGGGCACTGTGTGTGAAGAACAGACATGAGAGCTTTAGTGTGCACTTTGTGAT 2795
Db 2273 CGAAGAAATGGGCACTGTGTGTGAAGAACAGACATGAGAGCTTTAGTGTGCACTTTGTGAT 2432
Qy 2796 TAAATTCAGCCGCAACGTGAGAGAGCTTCAGCTGATGAGGGCAGGACAGATCAAC 2855
Db 2433 TAAATTCAGCCGCAACGTGAGAGAGCTTCAGCTGATGAGGGCAGGACAGATCAAC 2492
Qy 2856 ATGAGCCCCACACATGATAGTCTGTTCAGGTGGGTCCAGATCAAGATGCTATTGGCA 2915
Db 2493 ATGAGCCCCACACATGATAGTCTGTTCAGGTGGGTCCAGATCAAGATGCTATTGGCA 2552
Qy 2916 GATTCTCTTCCGCTCTCAAGATCAACAGAACTGAGAGGCTGAGACTTAAGTGA 2975
Db 2553 GATTCTCTTCCGCTCTCAAGATCAACAGAACTGAGAGGCTGAGACTTAAGTGA 2612
Qy 2976 CTGCTGAGCCACTGTGAGTGAAGAGTCTTTGAAGCCCTGAGAGCCCTGCTGCT 3035
Db 2613 CTGCTGAGCCACTGTGAGTGAAGAGTCTTTGAAGCCCTGAGAGCCCTGCTGCT 2672
Qy 3036 CCTGAGAACCTTGCGGTTGGCTGGCTGGCTGACAGCTGAGAGACTGCAAGACCTTGC 3095
Db 2673 CCTGAGAACCTTGCGGTTGGCTGGCTGGCTGACAGCTGAGAGACTGCAAGACCTTGC 2732
Qy 3096 CTTTGGGCTGAGAGCAACAGACCTTGCAGAGACTGAGAGGCTGAGACTTGAATGTCTAC 3155
Db 2733 CTTTGGGCTGAGAGCAACAGACCTTGCAGAGACTGAGAGGCTGAGACTTGAATGTCTAC 2792
Qy 3156 GGATGCTGAGAGCAACAGACCTTGCAGAGACTGAGAGGCTGAGAGGCTGAGAGGCT 3215
Db 2793 GGATGCTGAGAGCAACAGACCTTGCAGAGACTGAGAGGCTGAGAGGCTGAGAGGCT 2852
Qy 3216 ACTGAGCTGTGAGCTGTGAGCTGACAGCTGAGCTGAGCTGAGAGGCTGAGCTGTGCT 3275
Db 2853 ACTGAGCTGTGAGCTGTGAGCTGACAGCTGAGCTGAGCTGAGAGGCTGAGCTGTGCT 2912
Qy 3276 TAGTGCAGCCCCAGCTGAGAGAGCTAGACCTGAGAGAGAAACAACCTGATAGAGTTGG 3335
Db 2913 TAGTGCAGCCCCAGCTGAGAGAGCTAGACCTGAGAGAGAAACAACCTGATAGAGTTGG 2972
Qy 3336 CGTGGAGCTGCTGTGTGAGGGGCTCAGGCACTCGCCCTGCAAACTCAACCCCTGGGGCT 3395
Db 2973 CGTGGAGCTGCTGTGTGAGGGGCTCAGGCACTCGCCCTGCAAACTCAACCCCTGGGGCT 3032
Qy 3396 GGACCAAGCAACTCTGAGTGAATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3455
Db 3033 GGACCAAGCAACTCTGAGTGAATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3092
Qy 3456 ACCTCAGCTGTCTCATCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3515
Db 3093 GCCTCAGCTGTCTCATCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 3152
Qy 3516 GGATACGGGAGAGAGATGATATGACATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3575
Db 3153 GGATACGGGAGAGAGATGATATGACATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3212
Qy 3576 GAGGGCGGCTTCCCATGTTGTCTCAGAGCTAATCTCAAACTCCTGAGAGAGAGAGATCTT 3635
Db 3213 GAGGGCGGCTTCCCATGTTGTCTCAGAGCTAATCTCAAACTCCTGAGAGAGAGATCTT 3272

Qy 3636 CCGAATTCGTAATTTGAGAGAGAGAAAGCTCCCGAGAGTATACCGGAGAGACTTTGTG 3695
Db 3273 CCGAATTCGTAATTTGAGAGAGAGAAAGCTCCCGAGAGTATACCGGAGAGACTTTGTG 3332
Qy 3696 CGTGCCTTTCCTCCTCTCTCAAGGGAGACTGATACAGAGGCTTTGGGAGCTGAGTGA 3755
Db 3333 CGTGCCTTTCCTCCTCTCTCAAGGGAGACTGATACAGAGGCTTTGGGAGCTGAGTGA 3392
Qy 3756 CTTCTGGGGCCCAACGGGGCTGTGCTACTGAGTATGTCAGAAAGAAAGAACTGTGA 3815
Db 3393 CTTCTGGGGCCCAACGGGGCTGTGCTACTGAGTATGTCAGAAAGAAAGAACTGTGA 3452
Qy 3816 CCGAGTTCACTTCCCTGATGAGTGGCTCTCAACCGCTGGCCCAACAGGGTCTCTGCTTGT 3875
Db 3453 CCGAGTTCACTTCCCTGATGAGTGGCTCTCAACCGCTGGCCCAACAGGGTCTCTGCTTGT 3512
Qy 3876 GATAGAGAGAGCGGTGACCGTTGAGATGAAATTTGTGTGTGGAGACAGTTCTTGGGTGA 3935
Db 3513 GGTGAGAGAGAGCGGTGACCGTTGAGATGAAATTTGTGTGTGGAGACAGTTCTTGGGTGA 3572
Qy 3936 GATCAACCCACAGACAGCTGATGTGTGAGAGGCTCTGTGTGACATCAAGGCTGAGCC 3995
Db 3573 GATCAACCCACAGACAGCTGATGTGTGAGAGGCTCTGTGTGACATCAAGGCTGAGCC 3632
Qy 3996 TGGAGCTGTGAGAGCTGTGACCTTCCCTCACTTTGTGTGCTTCCAGGGGGCCATGTGGA 4055
Db 3633 TGGAGCGGTGAGAGCTGTGACCTTCCCTCACTTTGTGTGCTTCCAGGGGGCCATGTGGA 3692
Qy 4056 CACATCCCTGTTCCAAATGGGCGCACTTTAAAGAGAGAGAGTGTCTCTGAGAGAGCCAGC 4115
Db 3693 CACATCCCTGTTCCAAATGGGCGCACTTTAAAGAGAGAGAGTGTCTCTGAGAGAGCCAGC 3752
Qy 4116 CAGGTGAGAGCTGATCAATGTTGTGAAGAAACCCAGCTTCTCCCTTTGGAGTCT 4175
Db 3753 CAGGTGAGAGCTGATCAATGTTGTGAAGAAACCCAGCTTCTCCCTTTGGAGTCT 3812
Qy 4176 CTTGAAATATGATCCATTAATGCTCCCTGAGCTTCACTTCCCTGATGCTTGTGCTTTA 4235
Db 3813 CTTGAAATATGATCCATTAATGCTCCCTGAGCTTCACTTCCCTGATGCTTGTGCTTTA 3872
Qy 4236 CCACCGGCTCATCTGAGAGAGTCACTTCCACTTCACTGATCCCAATGAGCTGCTC 4295
Db 3873 CCACCGGCTCATCTGAGAGAGTCACTTCCACTTCACTGATCCCAATGAGCTGCTC 3932
Qy 4296 CATTGGAAGAGACTGAGCTGTGCTATCGAAGCCCTGAGAGAGACAGCTGTTCTGGA 4355
Db 3933 CATTGGAAGAGACTGAGCTGTGCTATCGAAGCCCTGAGAGAGACAGCTGTTCTGGA 3992
Qy 4356 GTTCTACGTTGGCCACTTTGGGATGAGGATCAAGCTGCAAGTGAAGAAAGATGA 4415
Db 3993 GTTCTACGTTGGCCACTTTGGGATGAGGATCAAGCTGCAAGTGAAGAAAGATGA 4052
Qy 4416 GACTCTGATGGAGGCTTGTGTGAAGACAGAGATCTCACTGCTGCACTTCTGAT 4475
Db 4053 GACTCTGATGGAGGCTTGTGTGAAGACAGAGATCTCACTGCTGCACTTCTGAT 4112
Qy 4476 CCTTCAGCCCGCATAGCCCTTCACTCTGAGTGCCTGAGAGAGTTGCTGCACTTTGT 4535
Db 4113 CCTTCAGCCCGCATAGCCCTTCACTCTGAGTGCCTGAGAGAGTTGCTGCACTTTGT 4172
Qy 4536 GGACCAAGTATGAGAGAGAGAGTGAAGCCCAAGTGAATCGGTGAGAGTTGCTTGGACAA 4595
Db 4173 GGACCAAGTATGAGAGAGAGAGTGAAGCCCAAGTGAATCGGTGAGAGTTGCTTGGACAA 4232
Qy 4596 ACTGCATGAGACAGGTGCTGAGCCAGAGACAGTACGAGAGGGTGTGCTGAGAGACAGAG 4655
Db 4233 ACTGCATGAGACAGGTGCTGAGCCAGAGACAGTACGAGAGGGTGTGCTGAGAGACAGAG 4292
Qy 4656 GCCCAGCCAGATGGAGAGCTGTTCACTTGAAGCCAGTCTTGGAGCCGGAAGTGCAGAA 4715
Db 4293 GCCCAGCCAGATGGAGAGCTGTTCACTTGAAGCCAGTCTTGGAGCCGGAAGTGCAGAA 4352

QY 4716 TGGACTCTACCAAGCCCTGGAAGAGACCCATCTCTCACTTATGGAAGCTCTGGAGAA 4775
 DB 4353 TGGACTCTACCAAGCCCTGGAAGAGACCCATCTCTCACTTATGGAAGCTCTGGAGAA 4412
 QY 4776 GGGCAGCAAAAAGAGAGCTCTGACCTGACAGAGCTGAGAGATCAACAGCCGCTTACCC 4835
 DB 4413 GGGCAGCAAAAAGAGAGCTCTGACCTGACAGAGCTGAGAGATCAACAGCCGCTTACCC 4472
 QY 4836 CTGAGAGCTCTGAGCTTGGCTGACCTCTCTTGGGCTCTCAAGTTCTTCTGCAAAAG 4895
 DB 4473 CTGAGAGCTCTGAGCTTGGCTGACCTCTCTTGGGCTCTCAAGTTCTTCTGCAAAAG 4532
 QY 4896 TTGGCAGCTGAGCTTGGCTGACCTCTCTTGGGCTCTCAAGTTCTTCTGCAAAAG 4955
 DB 4533 TTGGCAGCTGAGCTTGGCTGACCTCTCTTGGGCTCTCAAGTTCTTCTGCAAAAG 4592
 QY 4956 GCATTAATGCTGATGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5015
 DB 4593 GCATTAATGCTGATGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4652
 QY 5016 TCTCAGGAGATGCTGATGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5075
 DB 4653 TCTCAGGAGATGCTGATGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4712
 QY 5076 GGTGGCAG 5135
 DB 4713 GGTGGCAG 4772
 QY 5136 GAG 5195
 DB 4773 GAG 4832
 QY 5196 TCAGCGTGGAG 5255
 DB 4833 TCAGCGTGGAG 4892
 QY 5256 AACAG 5315
 DB 4893 AACAG 4952
 QY 5316 AATTGCTGAG 5375
 DB 4953 AATTGCTGAG 5012
 QY 5376 ACTTAATTTAAG 5435
 DB 5013 ACTTAATTTAAG 5072
 QY 5436 GTTA 5439
 DB 5073 GTTA 5076

RESULT 4
 AF298548 5100 bp mRNA linear PRI 19-FEB-2002
 LOCUS AF298548
 DEFINITION Homo sapiens caspase recruitment domain protein 7 mRNA, complete cds.
 VERSION AF298548.1 GI:10198208
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Bertin J. and Distefano P.S.
 TITLE The PYRIN domain: a novel motif found in apoptosis and inflammation proteins
 JOURNAL Cell Death Differ. 7 (12), 1273-1274 (2000)
 MEDLINE 21169419
 PUBMED 11270363
 REFERENCE 2 (bases 1 to 5100)

AUTHORS Bertin J.
 TITLE Direct Submission
 JOURNAL Submitted (23-AUG-2000) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA
 FEATURES
 source 1..5100
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /note="CARD; nucleotide-binding site/leucine-rich repeat family member (NBS/LRR); contains PYRIN and CARD domains"
 /codon_start=1
 /product="caspase recruitment domain protein 7"
 /protein_id="AA015254.1"
 /db_xref="GI:10198208"
 /translation="MAGAMGGLACYLEFLKEELKEPOLLLANKHSSSGEPAPQPKTSQEVAYSLVAOYGEORAMDLALHTWEMGRSLCAONDEGSHSPSPSEPHIGSPQPTSTAYVMPNHELPACTGSSERIVRGDLPTGRRWRBRSASHVALPSPHESPSQSPNAPSTAVLSWSPSPVLAPEBQAFQWPLDRTGITYTEIRERERKSEKRPMAAVGTPPOHSLQPHHPWESVSELSCTWPMNEDPQKFTOLLQRPHPRODPLVKSMDPYEENRGLIETIDLPGLDQEPRIVLOGAAGIKSTLAEQVKEAMGSGOLGDRFOHVFYSCRELAOSKVSIAELIGDGTAPAPIRILSRPRLFLIDGVDSPGWLQEPSELGLHNSQPOPADLGLSKTILPESPLITRTTLQNLISLEQARVETLGSSESRKRYFVPTLDEPQATAPFLPSKEELMALCLVFWWSMLACTCMQWRKREKRLTSLTTLTCLHTLALQAPLPGQURDLCSLAEGIWOKTLPSPDRLRKGDLGAILSTLRKGLIETLHPILSYSTLHCFQEFFAAMSIVLIDEBKRGKHSNCTIDLEKTLBAYGILHGLASTTRFLPLSDGEBREMNIPHCIRSNLMQWVPSLIQILQPLSLHCLYETRNTPFLQVMAHEE MGCWETDEMLVCTCFIKFSRHKVLLQIEBORHSTSPFMYVFRVAVPTDAYQILPSYLKYRNLKELTIGNSLSHVAVSLCTTLRPRCLTFLRLAGVLAEDKDLAFGLRANOTLLELDLSEVLMDAQKHLQRLRPSCTQLQLVSCSLTSDCCDLASVLASPLKELDLQNNLDVGVRLICEGLHAPCKLIRGLDITLSDMRQLALBOERKPOLIFSRKRSYMPDIBGLTEBMSNSTSLRORLGSBRASHVAQALKLIDVKEKIPPIAIAEESPEVPAVLCPSPASQGLDITKELGDDDFWGTGPVATBYVDEKMLYRHPVAGSVYRMPNGLCFVRAEAVTEIEFCVMDQFLIEINPDSMVAGPLDITKAEBAEVAHLPHFVALQGGHDTLSPOVHPEBGLKPARVEIHLTYLBNPSRPLDVLKMLHNLRIPLVTSVLYLHRLHEBETPHLYLIPSCSIRKELELCYRSPGEDLFSFVYGHLSGIRLVQVKOKKELTWEALVREGDIMPATLLPACIAVPSPLDAPQLHFDVQYREOLIRARTSVAVLADKLHGOVLQSOBYEYLAENSTRPQMKLFLISQMDRKCDGLYQALKEHPHILIMELKSGKGLPLPS"

BASE COUNT 1251 a 1374 c 1406 g 1069 t

ORIGIN

Query Match 89.7%; Score 4882.4; DB 9; Length 5100;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 516 GACAGAGATGCTGGGGAGAGCTGGGGCCGCTGGCTCTTCTGGAAGTTCCTGAAGA 575
 DB 153 GAACAAGATGGCTGGGGAGAGCTGGGGCCGCTGGCTCTTCTGGAAGTTCCTGAAGA 212
 QY 576 GAGAGAGCTGAAGAGATTCAGCTTCTGCTCCCATATAAGGCACTCCAGAGAGCTTTC 635
 DB 213 GAGAGAGCTGAAGAGATTCAGCTTCTGCTCCCATATAAGGCACTCCAGAGAGCTTTC 272
 QY 636 GAGTGAACACCCGCTCAGCCAGAGAAACGAGTGCATGAGAGTGGCTCTGTAAGTGGT 695
 DB 273 GAGTGAACACCCGCTCAGCCAGAGAAACGAGTGCATGAGAGTGGCTCTGTAAGTGGT 332
 QY 696 GGTCTAGTATGGGGAGACACGGGCTGGGACCTAGCCCTCCATACCTGGAGAGAGTGG 755
 DB 333 GGTCTAGTATGGGGAGACACGGGCTGGGACCTAGCCCTCCATACCTGGAGAGAGTGG 392
 QY 756 GCTGAGAGTACTGTGGCCCAAGCCAGAGAGAGGAGGCACTCTCCCTATTCCTCA 815
 DB 393 GCTGAGAGTACTGTGGCCCAAGCCAGAGAGAGGAGGCACTCTCCCTATTCCTCA 452
 QY 816 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGAGCAACCACTCCACCGCAGTGTAT 875
 DB 453 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGAGCAACCACTCCACCGCAGTGTAT 512
 QY 876 GCCCTGATTCATGAATTCGCGGCGGGGTGACCCAGAGGCTCAGAGAGAGAGGTTTGGAG 935

D	513	GCCCTGGATCATGATTTGGCGGGGGGTGACCCGAGGCTTCAAGAGAGAGTTTGAG	572
Q	936	ACAGCTGCTTGACATCTGGACGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCA	995
D	573	ACAGCTGCTTGACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTACACCTCTACCA	632
Q	996	AGCTCTTCCAAAGCTTCCCAAGACATAGTCTTCCAGCCAGAGATCCCAAGCCCCAC	1055
D	633	AGCTCTTCCAAAGCTTCCCAAGACATAGTCTTCCAGCCAGAGATCCCAAGCCCCAC	692
Q	1056	ATCCAGCAGAGTGTGGGGAGCTGGGGATGCCACCTCAGGCCAGGCTTACAGCCAGAGA	1115
D	693	ATCCAGCAGAGTGTGGGGAGCTGGGGATGCCACCTCAGGCCAGGCTTACAGCCAGAGA	752
Q	1116	GCAGAGAGCTCTGGAGACCCAGTGGCTCTGGATGAAACGTCAAGAAATTTTACACAGA	1175
D	753	GCAGAGAGCTCTGGAGACCCAGTGGCTCTGGATGAAACGTCAAGAAATTTTACACAGA	812
Q	1176	AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGAGCGGTGT	1235
D	813	AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGAGCGGTGT	872
Q	1236	AGGAAGCCCCCAAGGGGCAACAGGCTTACAGCCCAACCAACCATGGAGGCTTC	1295
D	873	AGGAAGCCCCCAAGGGGCAACAGGCTTACAGCCCAACCAACCATGGAGGCTTC	932
Q	1296	TGTGAGAGAGAGCTCTGTTCCACATGGGCTCTGAGAAATAGAGATTTTAAACAAAT	1355
D	933	TGTGAGAGAGAGCTCTGTTCCACATGGGCTCTGAGAAATAGAGATTTTAAACAAAT	992
Q	1356	CACACAGCTGTACTTCTACAAAGACCTCACCCAGAAAGCCAGATCCCTGTCAAAG	1415
D	993	CACACAGCTGTACTTCTACAAAGACCTCACCCAGAAAGCCAGATCCCTGTCAAAG	1052
Q	1416	AAGTGGCTGTATATGTGGAGAGAAATCCAGGACATTTAATTGAGATCAGAGCTTAAT	1475
D	1053	AAGTGGCTGTATATGTGGAGAGAAATCCAGGACATTTAATTGAGATCAGAGCTTAAT	1112
Q	1476	TGGCCAGAGCTGTGATACCCAAAGACCTCGATATGCTATCTGAGGGGGCTGTGGAT	1535
D	1113	TGGCCAGAGCTGTGATACCCAAAGACCTCGATATGCTATCTGAGGGGGCTGTGGAT	1172
Q	1536	TGGAGAGTCAACACTGGCCAGGAGGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTAG	1595
D	1173	TGGAGAGTCAACACTGGCCAGGAGGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTAG	1232
Q	1596	GGACCGCTTCCAGCATGTCTTCTACTTCACTGACGTGAGAGAGTGGCCCAAGTCAAGGTGT	1655
D	1233	GGACCGCTTCCAGCATGTCTTCTACTTCACTGACGTGAGAGAGTGGCCCAAGTCAAGGTGT	1292
Q	1656	GAGTCTGCTGAGCTCATCGGAAAGATGGGAGCCACTCGGCTCCCATTTAGACAT	1715
D	1293	GAGTCTGCTGAGCTCATCGGAAAGATGGGAGCCACTCGGCTCCCATTTAGACAT	1352
Q	1716	CCTGTCTTAGGCCAGAGCGGCTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGT	1775
D	1353	CCTGTCTTAGGCCAGAGCGGCTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGT	1412
Q	1776	CTTGCAAGAGCGAGTGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCGGGAGTGC	1835
D	1413	CTTGCAAGAGCGAGTGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCGGGAGTGC	1472
Q	1836	ACTGCTGGGGAGTTTGTGGGGGAAACATACTTCCGAGGCACTCTTCTGATCAGGCG	1895
D	1473	ACTGCTGGGGAGTTTGTGGGGGAAACATACTTCCGAGGCACTCTTCTGATCAGGCG	1532
Q	1896	TCGAGACACAGCTCTGAGAACTCTATTCTTCTTGGAGCAGGCAAGTTGGGTAGAGT	1955
D	1533	TCGAGACACAGCTCTGAGAACTCTATTCTTCTTGGAGCAGGCAAGTTGGGTAGAGT	1592
Q	1956	CTTGGGGTCTCTGAGTCCAGAGGAGAAATTTTCTACAGATATTTCAAGATGAAG	2015

QY	3096	CTTTGGGCTGAGAACCAACCAAGACCTTGACCCGAGCTGGACCTTGAGCTTCATATGCTCTAC	3155
Db	2733	CTTTGGGCTGAGAACCAACCAAGACCTTGACCCGAGCTGGACCTTGAGCTTCATATGCTCTAC	2792
QY	3156	GGATGCTGAGACCAACCACTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGTACAGCG	3215
Db	2793	GGATGCTGAGACCAACCACTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGTACAGCG	2852
QY	3216	ACTGCACTGCTGACGCTGTGGCTTCACGTCTGACTGCTGCCAGAGACTTGCCCTTGCTGCT	3275
Db	2853	ACTGCACTGCTGACGCTGTGGCTTCACGTCTGACTGCTGCCAGAGACTTGCCCTTGCTGCT	2912
QY	3276	TAGTGCAGAGCCCCGAGCTGAAGAGCTTACGCTGACGAGAAACAACCTGGATGACGTTGG	3335
Db	2913	TAGTGCAGAGCCCCGAGCTGAAGAGCTTACGCTGACGAGAAACAACCTGGATGACGTTGG	2972
QY	3336	CGTGCGACTGCTCTGTGAGGGGCTCAGGCATCTGCTGCCAATCATATACGCTTGCGCT	3395
Db	2973	CGTGCGACTGCTCTGTGAGGGGCTCAGGCATCTGCTGCCAATCATATAGCTTGCGGCT	3032
QY	3396	GGACCAACCAACTCTGAGTGTATGAGTACGACAGAACTGAGGGCCCTGAGACAGAGAA	3455
Db	3033	GGACCAACCAACTCTGAGTGTATGAGTACGACAGAACTGAGGGCCCTGAGACAGAGAA	3092
QY	3456	ACCTCAGCTGCTCATCTTCAGCAGACGGAAACAAGTGTATGACCCCTTACTGAGGGGCT	3515
Db	3093	GCTTCAGCTGCTCATCTTCAGCAGACGGAAACAAGTGTATGACCCCTTATGAGGGGCT	3152
QY	3516	GGATACCGGAGAGATGAGTAATAGCAATCTCTCAACGGCGGAGAGACTCGGATCAGA	3575
Db	3153	GGATACCGGAGAGATGAGTAATAGCAATCTCTCAACGGCGGAGAGACTCGGATCAGA	3212
QY	3576	GAGGGCGGCTTCCATGTTGTCTCAGGCGCTAACTCTGACCGAGACCAAGTCTT	3635
Db	3213	GAGGGCGGCTTCCATGTTGTCTCAGGCGCTAACTCTGACCGAGACCAAGTCTT	3272
QY	3636	CCCAATTGCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACTCTTGTC	3695
Db	3273	CCCAATTGCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACTCTTGTC	3332
QY	3696	CGTGCCTTCTCTGCTGCTCTCAAGGGGACCTGCAATCAAGGCTTTGGGGACTGACATGTA	3755
Db	3333	CGTGCCTTCTCTGCTGCTCTCAAGGGGACCTGCAATCAAGGCTTTGGGGACTGACATGTA	3392
QY	3756	CTTCTGGGGCCCCCAGGGGCTGTGGCTACCTGAGGTATGTACAAAGAAAGAACTTGTA	3815
Db	3393	CTTCTGGGGCCCCCAGGGGCTGTGGCTACCTGAGGTATGTACAAAGAAAGAACTTGTA	3452
QY	3816	CCGAGTTCACTTCCCTGTAGCTGACTCTACCGCTGGCCCAACAGGGTCTCTGCTTTGT	3875
Db	3453	CCGAGTTCACTTCCCTGTAGCTGACTCTACCGCTGGCCCAACAGGGTCTCTGCTTTGT	3512
QY	3876	GATAGAGGAAGCGGTGACCGTTGATGATTGATTCTGTGTGGGACCAAGTCTCTGGGGTA	3935
Db	3513	GATAGAGGAAGCGGTGACCGTTGATGATTGATTCTGTGTGGGACCAAGTCTCTGGGGTA	3572
QY	3936	GATCAACCCACAGCAGCAGCTGAGTGTGGCGGGCTCTGCTGACATCAAGGCTGAGCC	3995
Db	3573	GATCAACCCACAGCAGCAGCTGAGTGTGGCGGGCTCTGCTGACATCAAGGCTGAGCC	3632
QY	3996	TGAGAGCTGTGGAAGCTGTGCACCTTCCTCACTTTGTGCTCTCCAGGGGGCCATGTGGA	4055
Db	3633	TGAGAGCGGTGGAAGCTGTGCACCTTCCTCACTTTGTGCTCTCCAGGGGGCCATGTGGA	3692
QY	4056	CACATCCCTGTTCCAAAATGCGCCCACTTAAAGAGAGGGAGTGCCTCGAGAGAACCGC	4115
Db	3693	CACATCCCTGTTCCAAGTGGCCCACTTAAAGAGAGGGAGTGCCTCGAGAGAACCGC	3752
QY	4116	CAGGGTGAAGCTGCATCATATGTTCTGGAAAACCCACGTTCTCCCTTGGAGTCTT	4175
Db	3753	CAGGGTGAAGCTGCATCATATGTTCTGGAAAACCCACGTTTCTCCCTTGGAGTCTT	3812

QY	4176	CCGGAAGAAATGATCATTAATAGCCCTGGGCTCAATTCOCCGATCACTCGTGGGTGGCTTTA	4235
Db	3813	CTGAAATATATCATTAATAGCCCTGGGCTCAATTCOCCGATCACTCGTGGGTGGCTTTA	3872
QY	4236	CCACCCGCTCATCTTGAGGAAGTCACTTCCACCTTACTTGATCCAGATGACTGCTC	4295
Db	3873	CCACCCGCTCATCTTGAGGAAGTCACTTCCACCTTACTTGATCCAGATGACTGCTC	3932
QY	4296	CATTCCGAAGAACTGGAGCTCTGCTATGGAAGCCCTGGAGAAACCAAGCTGTTCTCGGA	4355
Db	3933	CATTCCGAAGAACTGGAGCTCTGCTATGGAAGCCCTGGAGAAACCAAGCTGTTCTCGGA	3992
QY	4356	GTCTTACGTTGGCCCATCTTGGAGATCAGGAGTCAAGGCTGCAAGTGAAGAACAAGAAATGA	4415
Db	3993	GTCTTACGTTGGCCCATCTTGGAGATCAGGAGTCAAGGCTGCAAGTGAAGAACAAGAAATGA	4052
QY	4416	GACTCTGTTGGAGAGCCCTTGGTGAACCAAGAGATCTCATGCTGCAACTACTTGAT	4475
Db	4053	GACTCTGTTGGAGAGCCCTTGGTGAACCAAGAGATCTCATGCTGCAACTACTTGAT	4112
QY	4476	CCCTCCAGCCCGGATAGCCGTAACCTTCACTCTGATGCCCCGAGTTGCTGCACTTGT	4535
Db	4113	CCCTCCAGCCCGGATAGCCGTAACCTTCACTCTGATGCCCCGAGTTGCTGCACTTGT	4172
QY	4536	GACACAGTATCGAGACAGCTGTATAGCCCGATGACATCGTGGAGGTTGTCTTGGACAA	4595
Db	4173	GACACAGTATCGAGACAGCTGTATAGCCCGATGACATCGTGGAGGTTGTCTTGGACAA	4232
QY	4596	ACTGCATGGAACAGGTGCTGATGCCAGAGCAATACAGAGGGTGTCTGGCTGAGAACACAG	4655
Db	4233	ACTGCATGGAACAGGTGCTGATGCCAGAGCAATACAGAGGGTGTCTGGCTGAGAACACAG	4292
QY	4656	GCCCAAGCAATCGGAAAGCTGTCAAGCTTGAAGCCAGTCTTGGAGCCGAAAGTGCAAGA	4715
Db	4293	GCCCAAGCAATCGGAAAGCTGTCAAGCTTGAAGCCAGTCTTGGAGCCGAAAGTGCAAGA	4352
QY	4716	TGACCTTACCAAGCCCTGAGAGAGCCCATCTCACTCATTTATGAACTCTGGAGAA	4775
Db	4353	TGACCTTACCAAGCCCTGAGAGAGCCCATCTCACTCATTTATGAACTCTGGAGAA	4412
QY	4776	GGGAGAGAAAGGAGCTGCTGCACTGACAGAGCTGAAGTATCAACAACAGCCCTTGACC	4835
Db	4413	GGGAGAGAAAGGAGCTGCTGCACTGACAGAGCTGAAGTATGAAACAACAGCCCTTGACC	4472
QY	4836	CTTGAGTCTTGAGCTTGGCTGACCCCTCTTGGGTCTCAATTCTTCTTCTGCAACAAG	4895
Db	4473	CTTGAGTCTTGAGCTTGGCTGACCCCTCTTGGGTCTCAATTCTTCTTCTGCAACAAG	4532
QY	4896	TTGCGATCTGTTGGCTTTCAGACATTAAGTAATGAACTTGAATGATGCTTGGCTGG	4955
Db	4533	TTGCGATCTGTTGGCTTTCAGACATTAAGTAATGAACTTGAATGATGCTTGGCTGG	4592
QY	4956	GCATTATGTGTCCATGTCAGAGGATGTCACAGAGGGGCCCATGTCAGAGTGGCTTAACACA	5015
Db	4593	GCATTATGTGTCCATGTCAGAGGATGTCACAGAGGGGCCCATGTCAGAGTGGCTTAACACA	4652
QY	5016	TCTCAGGGAATGTCCATCTTGAAGCTTGGAGAACCCCTGCAACTCATTAAGGCTCATCT	5075
Db	4653	TCTCAGGGAATGTCCATCTTGAAGCTTGGAGAACCCCTGCAACTCATTAAGGCTCATCT	4712
QY	5076	GGTGGCCACAGACAGCCAGCTTGAAGCCCTCGGAAATCCATCAGAGCGCAAAAGAGAAATA	5135
Db	4713	GGTGGCCACAGACAGCCAGCTTGAAGCCCTCGGAAATCCATCAGAGCGCAAAAGAGAAATA	4772
QY	5136	GGAAGGACATGAACCATTTGCTCTGGCTGTGTCAAGAGTGAAGCCCAAAATTGGGCT	5195
Db	4773	GGAAGGACATGAACCATTTGCTCTGGCTGTGTCAAGAGTGAAGCCCAAAATTGGGCT	4832
QY	5196	TGAGCGTGGGAGGCGACGTGGAATCTTGGCTTTGTAAGAGGAATCTCAAGAGCAAGCC	5255
Db	4833	TGAGCGTGGGAGGCGACGTGGAATCTTGGCTTTGTAAGAGGAATCTCAAGAGCAAGCC	4892
QY	5256	AACAGAGTAAAGTGAAGAAAGTTTATTCAGAAATTAAGAGATACAGGCTTTTAA	5315

Db 4893 AACGAGTAAGTGAAGAGAGTTTATTCAGAAATTAAGAGTATCTGCTCTTTAG 4952
Qy 5316 AATTGTCTAGCAGGCTTTCCAGTTTTCAGAAACCCCTATTAATTAATTTT 5375
Db 4953 AATTGTCTAGCAGGCTTTCCAGTTTTCAGAAACCCCTATTAATTAATTTT 5012
Qy 5376 ACTTAATTTTAAGATTTAAATAATCAAAAAGAAAAATGAATAATTAAGATTAAGA 5435
Db 5013 ACTTAATTTTAAGATTTAAATAATCAAAAAGAAAAATGAATAATTAAGATTAAGA 5072
Qy 5436 GTTA 5439
Db 5073 GTTA 5076

RESULT 5
AX207661

LOCUS AX207661 5122 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 29 from Patent WO0157085.
ACCESSION AX207661
VERSION AX207661.1 GI:15422343
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Baughn, M.R., Au-Young, J. and Yue, H.
TITLE G-protein coupled receptors
JOURNAL Patent: NO 0157085-A 29 09-AUG-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers

FEATURES
source 1..5122
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 4250893CB1"

BASE COUNT 1184 a 1460 c 1401 g 1077 t
ORIGIN

Query Match 87.5%; Score 4765.6; DB 6; Length 5122;

Best Local Similarity 96.9%; Pred. No. 0; Mismatches 24; Indels 133; Gaps 2;
Matches 4933; Conservative 0;

Qy 35 GAGCAGAGAGCCCGGAGCTCACTGAGTTCTGAAGCCATTCCTGCTGCGCTC 94
Db 1 GAGCAGAGAGCCCGGAGCTCACTGAGTTCTGAAGCCATTCCTGCTGCGCTC 60
Qy 95 CTCCACCCACCTCTTCTGAGCTTTCAGCTCAAGGTTGATCTCAGAGTCCAGAAC 154
Db 61 CTCCACCCACCTCTTCTGAGCTTTCAGCTCAAGGTTGATCTCAGAGTCCAGAAC 120
Qy 155 CAGAGAGAGAGAGATCTGAGAGACAGACAGAGAGGCTGCCACACCCCATCTCC 214
Db 121 CAGAGAGAGAGAGATCTGAGAGACAGACAGAGAGGCTGCCACACCCCATCTCC 180
Qy 215 GTCAACCAATCTCCCTCACTCACTCCCTGCTGCTGAGACCCCATCCAGAGAC 274
Db 181 GTCAACCAATCTCCCTCACTCACTCCCTGCTGCTGAGACCCCATCCAGAGAC 240
Qy 275 CTCCCTATCAGCTGATCTTCTCAAGTCTTTCAGAGGCCCCCTTGAGGCTCTCCCT 334
Db 241 CTCCCTATCAGCTGATCTTCTCAAGTCTTTCAGAGGCCCCCTTGAGGCTCTCCCT 300
Qy 335 GGGCTTTTCTACACATCCCTCTATCCGCGCTATCTGATAGTGCCCGGATTTATA 394
Db 301 GGGCTTTTCTACACATCCCTCTATCCGCGCTATCTGATAGTGCCCGGATTTATA 360
Qy 395 AACTGGGTTCCGAATGCTGAATTAAGAGAGGTAAGGCAAGCAAGCAAGCACTGTT 454
Db 361 AACTGGGTTCCGAATGCTGAATTAAGAGAGGTAAGGCAAGCAAGCAAGCACTGTT 420

Qy 455 CTCTGCGCTGATACCCCTACACACCTGGAGAAATCCCCAGACACCCCTTTAATCCG 514
Db 421 CTCTGCGCTGATACCCCTACACACCTGGAGAAATCCCCAGACACCCCTTTAATCCG 480
Qy 515 GAGACAGATGAGCTGAGAGAGCTTGGAGCCGCTGAGCTGTTACTTGGAGTTCGAGA 574
Db 481 GAGACAGATGAGCTGAGAGAGCTTGGAGCCGCTGAGCTGTTACTTGGAGTTCGAGA 540
Qy 575 AGAGAGAGCTGAAGAGAGTTCAGCTTCTGCTGCCAATTAAGCGCATTCAGAGAGCTTT 634
Db 541 AGAGAGAGCTGAAGAGAGTTCAGCTTCTGCTGCCAATTAAGCGCATTCAGAGAGCTTT 600
Qy 635 CGGCTGAGACACCCGCTTACAGCCAGAGAGAGAGAGAGTGGATGGAGTGGCTCTGATCC 694
Db 601 CGGCTGAGACACCCGCTTACAGCCAGAGAGAGAGAGTGGATGGAGTGGCTCTGATCC 660
Qy 695 TGGCTCAGTATGGAG 754
Db 661 TGGCTCAGTATGGAG 720
Qy 755 GGGTGAAGTCACTGTGCGCCCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 814
Db 721 GGGTGAAGTCACTGTGCGCCCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 815 ACAAGCCCAAGGAAACCCCACTGGAGTCTCCAGAGCAACCCCACTCCACCCGAGCTTA 874
Db 781 ACAAGCCCAAGGAAACCCCACTGGAGTCTCCAGAGCAACCCCACTCCACCCGAGCTTA 840
Qy 875 TGCCCTGATCCATGAATTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 934
Db 841 TGCCCTGATCCATGAATTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 935 GACAGCTGCTGACACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 994
Db 901 GACAGCTGCTGACACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 995 AAGCTCTTCCAGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1054
Db 961 AAGCTCTTCCAGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1055 CATCCACAG 1114
Db 1021 CATCCACAG 1080
Qy 1115 AGCAGAGAGCTCTGAG 1174
Db 1081 AGCAGAGAGCTCTGAG 1140
Qy 1175 AATTCAG 1234
Db 1141 AATTCAG 1200
Qy 1235 TAGGAG 1294
Db 1201 TAGGAG 1260
Qy 1295 CTGTGAGAGAGAGCTCTGTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1354
Db 1261 CTGTGAGAGAGAGCTCTGTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1355 TCACACAGCTGATCTTTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1414
Db 1321 TCACACAGCTGATCTTTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1415 GAAGCTGAGCTGATCTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1474
Db 1381 GAAGCTGAGCTGATCTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1475 TTGGCCAGAGCTTGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1534
Db 1441 TTGGCCAGAGCTTGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1535 TTGGAGATCAACACTGAGCAG 1594

Db 1501 TTGGGAAGTAAACATGGCCAGGAGGTGAAGAAAGCTGGGGAGAGGCACTGATG 1560
 Qy 1595 GGGACGCTTCCAGCATGTCTTCACTTCACTGACAGAGCTG9CCAGTCCAAAGTGG 1654
 Db 1561 GGGACGCTTCCAGCATGTCTTCACTTCACTGACAGAGCTG9CCAGTCCAAAGTGG 1620
 Qy 1655 TGAATCTCGGTGAGTCTCATCGGAAAGATGGAGACAGCACTCCGGCTCCCATTAAGACAG 1714
 Db 1621 TGAATCTCGGTGAGTCTCATCGGAAAGATGGAGACAGCACTCCGGCTCCCATTAAGACAG 1680
 Qy 1715 TCCGTCTAGGCGAGAGCGGCTGTCTTCACTTCACTGATGGTGAATGAGACAGAGTGGG 1774
 Db 1681 TCCGTCTAGGCGAGAGCGGCTGTCTTCACTTCACTGATGGTGAATGAGACAGAGTGGG 1740
 Qy 1775 TCTTGAAGAGCCGAGTGTGAGCTGTCTGCACTGAGACCCAGCCAGCCGCGGATG 1834
 Db 1741 TCTTGAAGAGCCGAGTGTGAGCTGTCTGCACTGAGACCCAGCCAGCCGCGGATG 1800
 Qy 1835 CACTGCTGGGAGTGTGCTGGGGAAAATATATCTTCCGAGGCACTCTTCTGATCA 1894
 Db 1801 CACTGCTGGGAGTGTGCTGGGGAAAATATATCTTCCGAGGCACTCTTCTGATCA 1860
 Qy 1895 CTGGAGCAAGCTCTGAGCAAGCTCACTTCTTGGAGAGGCACTTGGGTAGAG 1954
 Db 1861 CTGGAGCAAGCTCTGAGCAAGCTCACTTCTTGGAGAGGCACTTGGGTAGAG 1920
 Qy 1955 TCTTGGGGTCTCTGAGTCCAGCAGAGAGATATTTCTACAGATATTTCAAGATGAAA 2014
 Db 1921 TCTTGGGGTCTCTGAGTCCAGCAGAGAGATATTTCTACAGATATTTCAAGATGAAA 1980
 Qy 2015 GCGAAGCAATTAAGCTTTAGTGTGTCATTAATCAAAAGAGCTTGGGCTGTGTC 2074
 Db 1981 GCGAAGCAATTAAGCTTTAGTGTGTCATTAATCAAAAGAGCTTGGGCTGTGTC 2040
 Qy 2075 TTGTGCTGGGGTCTGCTGGCTGCGCTGCACTTGGCTGAGCAGAGTGAACGGAGG 2134
 Db 2041 TTGTGCTGGGGTCTGCTGGCTGCGCTGCACTTGGCTGAGCAGAGTGAACGGAGG 2100
 Qy 2135 AAAAATCACTGATCTTCCAGACACACCAAACTCTGTCTTCACTTACCTTGGCCAGG 2194
 Db 2101 AAAAATCACTGATCTTCCAGACACACCAAACTCTGTCTTCACTTACCTTGGCCAGG 2160
 Qy 2195 CTCTCCAGCTCAGCATTGGAGACCCAGCTCAGAGACCTCTGCTCTGGCTGAGG 2254
 Db 2161 CTCTCCAGCTCAGCATTGGAGACCCAGCTCAGAGACCTCTGCTCTGGCTGAGG 2220
 Qy 2255 GCATCTGGGCAAAAAGACCTTTTCACTGATGATGCTCAGAGACATGGGTAGATG 2314
 Db 2221 GCATCTGGGCAAAAAGACCTTTTCACTGATGATGCTCAGAGACATGGGTAGATG 2280
 Qy 2315 GGGCCATCATCTCCACTTCTTGAAGATGGGTATTTCTTCAAGAGACCCCATCCCTGTA 2374
 Db 2281 GGGCCATCATCTCCACTTCTTGAAGATGGGTATTTCTTCAAGAGACCCCATCCCTGTA 2340
 Qy 2375 GCTACAGCTTCACTTCACTTCTTCAAGATGTTCTTGAAGACATGTCTTATGTCTGG 2434
 Db 2341 GCTACAGCTTCACTTCACTTCTTCAAGATGTTCTTGAAGACATGTCTTATGTCTGG 2400
 Qy 2435 AGATGAGAGGGAGAGTAAATCTAATGACATGATGATGATGATGATGATGATGATGATG 2494
 Db 2401 AGATGAGAGGGAGAGTAAATCTAATGACATGATGATGATGATGATGATGATGATGATG 2460
 Qy 2495 AAGCATATGAAATACATGAGCTGTTGGGAGATCAACACGTTTCTATTTGGGCTGT 2554
 Db 2461 AAGCATATGAAATACATGAGCTGTTGGGAGATCAACACGTTTCTATTTGGGCTGT 2520
 Qy 2555 TAAATGATGAGGGAGAGAGATGAGAAATCTTCACTGCTGGCTGTCTCAAGGGA 2614
 Db 2521 TAAATGATGAGGGAGAGAGATGAGAAATCTTCACTGCTGGCTGTCTCAAGGGA 2580
 Qy 2615 GGAACCTGATGAGTGGTCCCGTCCGAGAGCTGCTGACAGCACTCTGAGT 2674

Db 2581 GGAACCTGATGAGTGGTCCCGTCCCTGACAGTCTGCTGACAGCACTCTGAGT 2640
 Qy 2675 CCTCCATGCTTTGATGAGACTCGAGACAAAAGTCTCTGACACAAAGTATGCCAT 2734
 Db 2641 CCTCCATGCTTTGATGAGACTCGAGACAAAAGTCTCTGACACAAAGTATGCCAT 2700
 Qy 2735 TCGAAGAAATGGGATGTGTGTGAAGACAGACATGAGCTTTATGTGTGACTTTCTGCA 2794
 Db 2701 TCGAAGAAATGGGATGTGTGTGAAGACAGACATGAGCTTTATGTGTGACTTTCTGCA 2760
 Qy 2795 TTAATTCAGCCGCCACGTAAGAGCTTCACTGATGATGAGGACAGACAGATCA 2854
 Db 2761 TTAATTCAGCCGCCACGTAAGAGCTTCACTGATGATGAGGACAGACAGATCA 2820
 Qy 2855 CATGAGCCCAACATGATGATCTGTTGAGGTGGTCCAGTCAAGATGCTATTTGGC 2914
 Db 2821 CATGAGCCCAACATGATGATCTGTTGAGGTGGTCCAGTCAAGATGCTATTTGGC 2880
 Qy 2915 AGATTTCTTTCTCGCTCTCAAGGTCAAGAACTTGAAGAGCTGACCTTAATGGA 2974
 Db 2881 AGATTTCTTTCTCGCTCTCAAGGTCAAGAACTTGAAGAGCTGACCTTAATGGA 2940
 Qy 2975 ACTGCTGAGCACTGAGTGAAGAGCTTTGTAAGACCTTGAAGAGCTGAGCTGAGC 3034
 Db 2941 ACTGCTGAGCACTGAGTGAAGAGCTTTGTAAGACCTTGAAGAGCTGAGCTGAGC 3000
 Qy 3035 TCTGAGAGACCTTGGAGTGGCTGGCTGAGCTTCACTGAGAGCTGAGAGCTTGG 3094
 Db 3001 TCTGAGAGACCTTGGAGTGGCTGGCTGAGCTTCACTGAGAGCTGAGAGCTTGG 3060
 Qy 3095 CTTTGGGCTGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 3154
 Db 3061 CTTTGGGCTGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 3120
 Qy 3155 CGGATCTGAGAGCAAACTTTGACAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 3214
 Db 3121 CGGATCTGAGAGCAAACTTTGACAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 3180
 Qy 3215 GACTGAGCTGTGACAGCTGTGACCTGACCTGACCTGACCTGACCTGACCTGAC 3274
 Db 3181 GACTGAGCTGTGACAGCTGTGACCTGACCTGACCTGACCTGACCTGACCTGAC 3240
 Qy 3275 TTATGCTGAGCCCTGAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 3334
 Db 3241 TTATGCTGAGCCCTGAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 3300
 Qy 3335 GGGTGGAGCTGTGAGTGGGGCTCAGGACCTGAGCTGAGCTGAGCTGAGCTGAG 3394
 Db 3301 GGGTGGAGCTGTGAGTGGGGCTCAGGACCTGAGCTGAGCTGAGCTGAGCTGAG 3260
 Qy 3395 TGAACAGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3454
 Db 3361 TGAACAGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
 Qy 3455 AACCTGAGCTGATCTTCAAGACAGAGAACTGATGATGATGATGATGATGATGATG 3514
 Db 3421 AACCTGAGCTGATCTTCAAGACAGAGAACTGATGATGATGATGATGATGATGATG 3480
 Qy 3515 TGAATAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3574
 Db 3481 TGAATAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3540
 Qy 3575 AGAGGGCGGCTCCATGTTGCTCAGGCTAATCTCAACTCTGAGAGTGAAGATCT 3634
 Db 3541 AGAGGGCGGCTCCATGTTGCTCAGGCTAATCTCAACTCTGAGAGTGAAGATCT 3600
 Qy 3635 TCCCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3694
 Db 3601 TCCCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3660
 Qy 3695 GCGTGCCTTCTCTGCTCTCAAGGGGACTGATCAAGAGCTTTGGGGACTGAGATG 3754
 Db 3661 GCGTGCCTTCTCTGCTCTCAAGGGGACTGATCAAGAGCTTTGGGGACTGAGATG 3720

QY 3755 ACTTGGGGGGCCCAAGGGGCTGTGGTACTGAGTATGTTGACAAAGAAAGAACTTGT 3814
 DB 3721 ACTTGGGGGGCCCAAGGGGCTGTGGTACTGAGTATGTTGACAAAGAAAGAACTTGT 3780
 QY 3815 ACCGAGTTCACTTCCCTGTAGTGGCTCTTACCGCTGAGCCCAACAGGGGTCTTGTCTTG 3874
 DB 3781 ACCGAGTTCACTTCCCTGTAGTGGCTCTTACCGCTGAGCCCAACAGGGGTCTTGTCTTG 3840
 QY 3875 TGATGAGAGAGCGGTGACCGTTGAGATTAATCTGTGTGTGGAGCCAGTCTTGGGTG 3934
 DB 3841 TGATGAGAGAGCGGTGACCGTTGAGATTAATCTGTGTGTGGAGCCAGTCTTGGGTG 3900
 QY 3935 AGATCAACCCACAGCAGCTGTGATGTGTGAGAGGGCTCTGTGACATCAAGCTGAGC 3994
 DB 3901 AGATCAACCCACAGCAGCTGTGATGTGTGAGAGGGCTCTGTGACATCAAGCTGAGC 3960
 QY 3995 CTGAGAGCTGTGAGAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGAGGGGGCCATGTGG 4054
 DB 3961 CTGAGAGCTGTGAGAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGAGGGGGCCATGTGG 4020
 QY 4055 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGAGTCTCTGAGAGAGCCAG 4114
 DB 4021 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGAGTCTCTGAGAGAGCCAG 4080
 QY 4115 CCAGGGTGTGAGCTGTGATCACTAATGTTCTGGAAGAAACCCAGCTTCTCCCTTGGAGTCC 4174
 DB 4081 CCAGGGTGTGAGCTGTGATCACTAATGTTCTGGAAGAAACCCAGCTTCTCCCTTGGAGTCC 4140
 QY 4175 TCCTGAAGAAATGATCATTAATGCTCCCTGAGCTCATTTCCGCTGACCTGTGTGTGTGCTT 4234
 DB 4141 TCCTGAAGAAATGATCATTAATGCTCCCTGAGCTCATTTCCGCTGACCTGTGTGTGTGCTT 4200
 QY 4235 ACCACCGGTCCATCTGTGAGAGAGTCACTTTCACCTCTTACCTGATCCCAAGTACTGCT 4294
 DB 4201 ACCACCGGTCCATCTGTGAGAGAGTCACTTTCACCTCTTACCTGATCCCAAGTACTGCT 4260
 QY 4295 CCATTTCG----- 4302
 DB 4261 CCATTTCGAGAGGCCATGATGATCTAGAAATGAAATTCAGTTTGTGGAAATCCAGAC 4320
 QY 4303 ----- 4302
 DB 4321 CACCCCGGCTGACCCCACTTTAATATGGGCTCTGTTACACTGTGTCTGGGTCTGCTTACG 4380
 QY 4303 -----AAGAACTGAGACTCTGTCTATCGAAGCCCTGAGAAAGACC 4342
 DB 4381 GGATGCTGAAATACTCCCAAGGAATGGAGCTCTGTCTATCGAAGCCCTGAGAAAGACC 4440
 QY 4343 AGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCAAGGATCAGGCTGCAAGTGAAG 4402
 DB 4441 AGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCAAGGATCAGGCTGCAAGTGAAG 4500
 QY 4403 ACAAGAAAGATGAGACTGTGTGTGGAGGCTTGGTGAACCAAGAGATCTCATGCTGTG 4462
 DB 4501 ACAAGAAAGATGAGACTGTGTGTGGAGGCTTGGTGAACCAAGAGATCTCATGCTGTG 4560
 QY 4463 CAACTACTCTGATCCCTTCAGCCCGCATAGCCGATACCTTCACTCTGTGATCCCGCAGT 4522
 DB 4561 CAACTACTCTGATCCCTTCAGCCCGCATAGCCGATACCTTCACTCTGTGATCCCGCAGT 4620
 QY 4523 TGCTGCACTTTGTGGAACAATGAGAGAGCTGTATGAGCCGAGTGAATCGGTGGAGG 4582
 DB 4621 TGCTGCACTTTGTGGAACAATGAGAGAGCTGTATGAGCCGAGTGAATCGGTGGAGG 4680
 QY 4583 TTGCTTTGGAACAATGAGAGAGCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4642
 DB 4681 TTGCTTTGGAACAATGAGAGAGCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
 QY 4643 CTGAGAAACAGAGGCCAGCCAGATGCGAAAGCTGTTCAAGTTGAGCCAGTCTGGAGAC 4702
 DB 4741 CTGAGAAACAGAGGCCAGCCAGATGCGAAAGCTGTTCAAGTTGAGCCAGTCTGGAGAC 4800

QY 4703 GGAAGTCAAGATGACTCTTACCAAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4762
 DB 4801 GGAAGTCAAGATGACTCTTACCAAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
 QY 4763 AACTCTGGAG 4822
 DB 4861 AACTCTGGAG 4920
 QY 4823 CCAGCCCTTGAACCTTGAATCTGTGCTTGTGCTGAGTCACTTCTTGTGGTCTTCAATTTCTT 4882
 DB 4921 CCAGCCCTTGAACCTTGAATCTGTGCTTGTGCTGAGTCACTTCTTGTGGTCTTCAATTTCTT 4980
 QY 4883 CTCTGCAAAACAAGTGGCATCTGTGTTGCTTCCAGACATTAAGTAATGAGAACTTATG 4942
 DB 4981 CTCTGCAAAACAAGTGGCATCTGTGTTGCTTCCAGACATTAAGTAATGAGAACTTATGAT 5040
 QY 4943 ATGCTTGTGCTGGGCACTTATGATGTCATGTCAGAGGAGATGCCACAGGGGGCC 4992
 DB 5041 GATGCTTGTGCTGGGCACTTATGATGTCATGTCAGAGGAGATGCCACAGGGGGCC 5089

RESULT 6
 AF229059 4290 bp mRNA linear PRI 02-FEB-2001
 LOCUS Homo sapiens NAC-alpha splice variant (NAC) mRNA, complete cds,
 DEFINITION alternatively spliced.
 ACCESSION AF229059
 VERSION AF229059.1 GI:12656104
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTORS Chu, Z.-L., Pio, F., Xie, Z., Godzik, A., and Reed, J.C.
 TITLE NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c
 pathway for apoptosis

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4290)
 AUTHORS Chu, Z.-L., and Reed, J.C.
 TITL Direct Submission
 JOURNAL Submitted (28-JUN-2000) Apoptosis and Signal Transduction, The
 Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA
 92037, USA

FEATURES

source

Location/Qualifiers

1..4290

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="11"

/cell_type="activated Jurkat T cells"

1..4290

/gene="NAC"

1..4290

/gene="NAC"

/note="Apaf-1/Ced-4 family apoptosis regulator; contains
NB-ARC domain, CARD domain, and leucine-rich repeats
(LRR)"

/codon_start=1

/product="NAC-alpha splice variant"

/protein_id="AAK00748.1"

/db_xref="GI:12656105"

/translation="MAGGAGRLACYLEFLKEKELEFOLLANKRAHSSSGETPAO
PEKTSQMEVASYLVAOYGBQAWDLHTWEOMGLRLCAOQEGAGHSPSPYSPE
PHLSPSPSTAVLWPMWIELPAGCTGSEERVLRLPDTGSRMRPEISALYLQAL
PSBPDSPEQSPNAPSTAVLWPMWIELPAGCTGSEERVLRLPDTGSRMRPEISALYLQAL
IRREKREKSRGPRMAAVVCTPQAAHSLQPHNHPERPVRSLSGTWKNEDDYE
KTPOLLQRPNRSQDPLVKSMPRVLENKHLIETRLRFGGLDTQPRVYLIG
AAGIGSTLARKQVKEAWRGQLYGNRPHVFLSCRLAQSVAVALIELIGKDTATP
APRQLISRPERLFLIDVDEPQWVLQGESSSLCLWSPQPADALGLSLKTTIP
EASFLIATRTALONLIPSLQARWVEVLGESSSRKEYRYRFTDERQAIRFLVYK
SNREMLALCVPAVSWLACTCLMOMKREKLTLTSTTTTLCILHYLAQLOAPLP
QARDLSIAAGIWMQKTLFSPDRLRHGIDGAIISFTLWGLIQEHPILSYFHL

CRQEFRAAMSVYLEDKRGKSHNCIIDEKTLFAYGHGFGASTTRFELGLSDEG
EREMENIPIHCRISQGRNLMOWPSLQILLDOPHSLSHCLYETRNTRFTLQVNAHRE
MGWCVETDEMLIVCFICIKFSRHVKLQILBGRQHSRWSPTMVLFRWPVDAWQ
LLEFVLKVRNMLKELDSGNSLSHSAVSKLRLRPRCLLETARLSCGLTAECDL
LAEGLRANOTLELDELSDNLSHFNLDAGAHLCQRLRPSCKLRLQVSCGLTAECDL
ASVLSASPILKELDLQONNLDVGVRLGGLHNPACILRIGLDOTTLSDEMRGLR
ALBOKRDLIFSRKRDQVMTPEGLDTEMMSNSTSLRQKDETLVMEALVKGDLMPATLLP
LIDVSKITPIAEIAEBSPREVVELCTPSPASQDHTKPIGTDDDRWGPYPAT
EVDKERNLRYHFPVAGSTRMNTGLCFVREAVLVELEPCWMDPFLBINFQHSWM
VAGPLDIDKABPQAVAVAHLPFVALQGVHVDLSFOMAFKEBGMLEBPARELHM
IYLENSFSPGLVLLKMHNLARFIPTVSVLLHYHVEVPEVLYLIPSDSIRKE
LEICTYSPGEBDQFSEFPYVGLSGAIRLOVKDKDETLVMEALVKGDLMPATLLP
ARIAPVSPDLAPOLHFDVQYREBOLJARTVSVVULDKLHGQVLSQBOYERVLAEHTR
PSQMKLFSLSQSWDRKCKDGLYQALKETHPHLIMEWKGSKGLPLPSS"

BASE COUNT 1011 a 1189 c 1209 g 881 t

ORIGIN

Query Match 78.8%; Score 4290; DB 9; Length 4290;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 ATGGCTGGCGAGCCTGGGGCCGCTGGCTGTACTTGAAGTTCCTGAAGAAGAGAG 582
Db 1 ATGGCTGGCGAGCCTGGGGCCGCTGGCTGTACTTGAAGTTCCTGAAGAAGAGAG 60
QY 583 CTGAAGAAGTTCAGCTTCTGCTGCGCAATAAAGCGCATCTCAGAGACTTTGGGGTAG 642
Db 61 CTGAAGAAGTTCAGCTTCTGCTGCGCAATAAAGCGCATCTCAGAGACTTTGGGGTAG 120
QY 643 ACAACCGCTCAGCAGAGAAGACGAGTGGCATGGAAGGTGGCTGTACTTGGGGTAG 702
Db 121 ACAACCGCTCAGCAGAGAAGACGAGTGGCATGGAAGGTGGCTGTACTTGGGGTAG 180
QY 703 TATGGGAGCAGCGGGCTGGGAGCTAGCCCTCATACCTGGAGACAGATGGGGCTGAGG 762
Db 181 TATGGGAGCAGCGGGCTGGGAGCTAGCCCTCATACCTGGAGACAGATGGGGCTGAGG 240
QY 763 TCACTGGGCGCCAAAGCCCAAGAAAGGGGAGGCGCATCTCTCCCTACATCCCTTACAGCCCA 822
Db 241 TCACTGGGCGCCAAAGCCCAAGAAAGGGGAGGCGCATCTCTCCCTACATCCCTTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGGTCTCCAGAGCAACCACTCCAGCGAGTGAATGGCCCTGG 882
Db 301 AGTGAACCCCACTGGGGGTCTCCAGAGCAACCACTCCAGCGAGTGAATGGCCCTGG 360
QY 883 ATCCATGAATTCGCGGCGGGGTGACCCAGAGGCTCAGAGAAAGGTTTGAAGACGCTG 942
Db 361 ATCCATGAATTCGCGGCGGGGTGACCCAGAGGCTCAGAGAAAGGTTTGAAGACGCTG 420
QY 943 CCTGAACATCTGAGACGCGCTGAGAGAAATCTTCTGCTCATCTCTTACCAAGCTCTT 1002
Db 421 CCTGAACATCTGAGACGCGCTGAGAGAAATCTTCTGCTCATCTCTTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAAGACCATAGTCTCCAAAGCGAGAGTCAACCAAGCGCCCAATCCACA 1062
Db 481 CCAAGCTCCCAAGACCATAGTCTCCAAAGCGAGAGTCAACCAAGCGCCCAATCCACA 540
QY 1063 GCAAGTGTGGGAGCTGGGGATCCCACTCAGGCCAGCTTAGCACCAGAGACGAGAG 1122
Db 541 GCAAGTGTGGGAGCTGGGGATCCCACTCAGGCCAGCTTAGCACCAGAGACGAGAG 600
QY 1123 GCTCCCTGGGACCCCAATGGCTCTTGGATGAAAGCTCAGGAATTTTACTACAGAAATCAGA 1182
Db 601 GCTCCCTGGGACCCCAATGGCTCTTGGATGAAAGCTCAGGAATTTTACTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCAGAGCCCACTGGGAGCGGTGTAGAGACG 1242
Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGCAGAGCCCACTGGGAGCGGTGTAGAGACG 720
QY 1243 CCCCCAAGGCGCACACAGCTTACAGCCCAACCAACCAATGGAGGCTTCTGTAGAG 1302
Db 721 CCCCCAAGGCGCACACAGCTTACAGCCCAACCAACCAATGGAGGCTTCTGTAGAG 780

QY 1303 GAGAGCTCTGTTCACATGGGCCCTGAAAAAATGAGAGATTTTAAACCAAAATTCACACAG 1362
Db 781 GAGAGCTCTGTTCACATGGGCCCTGAAAAAATGAGAGATTTTAAACCAAAATTCACACAG 840
QY 1363 CTGCTACTCTTCAAAAGACCTCAGCCACAGAGCAAGATCCCTGGTCAAGAGAACTGG 1422
Db 841 CTGCTACTCTTCAAAAGACCTCAGCCACAGAGCAAGATCCCTGGTCAAGAGAACTGG 900
QY 1423 CCTGATTATGTGAGAGAGATCGAGACATTTAATTGAGATCAAGACTTATTTGGCCCA 1482
Db 901 CCTGATTATGTGAGAGAGATCGAGACATTTAATTGAGATCAAGACTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCCCAAGAACCTGACATAGTCATATCTGACAGGGGGCTGCTGAAATTTGGAGAG 1542
Db 961 GGCCTGGATACCCCAAGAACCTGACATAGTCATATCTGACAGGGGGCTGCTGAAATTTGGAGAG 1020
QY 1543 TCACACATGAGCAGGAGTGAAGAGAGCCTGGGGGAGAGCCAGCTGTATGGAGACCGC 1602
Db 1021 TCACACATGAGCAGGAGTGAAGAGAGCCTGGGGGAGAGCCAGCTGTATGGAGACCGC 1080
QY 1603 TTCCAGCATGTCTTCTTACTTACAGCTGACAGAGCTGGCCAGTCCAAAGGTGTAGTCTTC 1662
Db 1081 TTCCAGCATGTCTTCTTACTTACAGCTGACAGAGCTGGCCAGTCCAAAGGTGTAGTCTTC 1140
QY 1663 GCTGAGCTATGAGAAAGATGGGACAGCCACTCCGGCTCCCAATTGACAGATCCGTCT 1722
Db 1141 GCTGAGCTATGAGAAAGATGGGACAGCCACTCCGGCTCCCAATTGACAGATCCGTCT 1200
QY 1723 AGGCGAGAGCGGCTGCTTCTTACCTCGATGGTGTATGATGAGACCAAGATGGGTCTTGACAG 1782
Db 1201 AGGCGAGAGCGGCTGCTTCTTACCTCGATGGTGTATGATGAGACCAAGATGGGTCTTGACAG 1260
QY 1783 GAGCGAGTTCGAGCTCTGTCTGCACTGAGGCGACCCAGCGCGGAGTGAATCTGCTG 1842
Db 1261 GAGCGAGTTCGAGCTCTGTCTGCACTGAGGCGACCCAGCGCGGAGTGAATCTGCTG 1320
QY 1843 GCGAGTTGCTGGGGAAACTATACCTCCGAGGACCTCTTCTGATCAAGGCTGGAGCC 1902
Db 1321 GCGAGTTGCTGGGGAAACTATACCTCCGAGGACCTCTTCTGATCAAGGCTGGAGCC 1380
QY 1903 ACAGCTCTGACAGAACCTCATCTCTTCTTGGAGACAGGACGTTGGGTAGAGGCTCTGGGG 1962
Db 1381 ACAGCTCTGACAGAACCTCATCTCTTCTTGGAGACAGGACGTTGGGTAGAGGCTCTGGGG 1440
QY 1963 TTCTCTGAGTCCAGCAGAGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGAGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 1500
QY 2023 ATTGAAGCCTTTAGTGGTGCATTAACAACAAGAGCTCGGGCCCTGTGCTGTGGGCC 2082
Db 1501 ATTGAAGCCTTTAGTGGTGCATTAACAACAAGAGCTCGGGCCCTGTGCTGTGGGCC 1560
QY 2083 TGGGTGCTCTGGCTGGCTGACCTTGCCTGATGACAGATGAAAGGAAAGAAAACTC 2142
Db 1561 TGGGTGCTCTGGCTGGCTGACCTTGCCTGATGACAGATGAAAGGAAAGAAAACTC 1620
QY 2143 ACACTGACTTCAAGACACCAACCACTCTGTCTTACATTAACCTTGCACAGGCTCTCCAA 2202
Db 1621 ACACTGACTTCAAGACACCAACCACTCTGTCTTACATTAACCTTGCACAGGCTCTCCAA 1680
QY 2203 GCTCAGCCATTTGGGACCCCAAGCTCAGAGACCTCTGCTCTGGCTGAGAGGCAATCTGG 2262
Db 1681 GCTCAGCCATTTGGGACCCCAAGCTCAGAGACCTCTGCTCTGGCTGAGAGGCAATCTGG 1740
QY 2263 CAAAAAAGAACCTTTTCACTCAGATGACCTCAGAGACATGGGTTAGATGGGCGCATC 2322
Db 1741 CAAAAAAGAACCTTTTCACTCAGATGACCTCAGAGACATGGGTTAGATGGGCGCATC 1800
QY 2323 ATCTCACTCTTCTTGAAGATGGGTATTTTCAAGAGACACCCATCCCTCTGAGCTACAGC 2382
Db 1801 ATCTCACTCTTCTTGAAGATGGGTATTTTCAAGAGACACCCATCCCTCTGAGCTACAGC 1860
QY 2383 TTCAATTCACCTGTTTCAAGAGTTCCTTGGAGCAATCTCATATGCTTGGAGATGAG 2442

Db 1861 TTCAATCACTCTGTTTCCAAAGATTCTTTGACGAATGCTTATGTTTGAGAGATGAG 1920
Qy 2443 AAGGGAGAGGTAAACATTCTAATTGCATCATAGATTGAAAGACGCTAGAACATAT 2502
Db 1921 AAGGGAGAGGTAAACATTCTAATTGCATCATAGATTGAAAGACGCTAGAACATAT 1980
Qy 2503 GGAATACATGAGCCCTGTTTGGGGCATCAACACGTTTCTTAATTGGGCTGTTAAGTAT 2562
Db 1981 GGAATACATGAGCCCTGTTTGGGGCATCAACACGTTTCTTAATTGGGCTGTTAAGTAT 2040
Qy 2563 GAGGGGAGAGAGATGAGAGACATCTTTCACTGCGCGCTGTCTCAGGGAGAGAACCTG 2622
Db 2041 GAGGGGAGAGAGATGAGAGACATCTTTCACTGCGCGCTGTCTCAGGGAGAGAACCTG 2100
Qy 2623 ATGCAATGAGATCCCGTCCCTGCAAGCTGCTGTCAGCCACATCTCTGAGATCCCTCAC 2682
Db 2101 ATGCAATGAGATCCCGTCCCTGCAAGCTGCTGTCAGCCACATCTCTGAGATCCCTCAC 2160
Qy 2683 TGCTTTGACAGATCTCGGACAAAAGTTCTTGACACAGATGATGGCCATTGAGAGAA 2742
Db 2161 TGCTTTGACAGATCTCGGACAAAAGTTCTTGACACAGATGATGGCCATTGAGAGAA 2220
Qy 2743 ATGGGCAATGATGATGAGAAACAGACATGGAGCTTATGATGCACTTCTGATTAATTC 2802
Db 2221 ATGGGCAATGATGATGAGAAACAGACATGGAGCTTATGATGCACTTCTGATTAATTC 2280
Qy 2803 AGCCGCCACGTGAGAGACCTTCAAGCTGATGAGGCGAGGACACAGATCAACATGAGC 2862
Db 2281 AGCCGCCACGTGAGAGACCTTCAAGCTGATGAGGCGAGGACACAGATCAACATGAGC 2340
Qy 2863 CCCACCATGATGATCTGTTTCAAGTGGGTCCCATGTCACAGATGCTTATGCAAGATTCTC 2922
Db 2341 CCCACCATGATGATCTGTTTCAAGTGGGTCCCATGTCACAGATGCTTATGCAAGATTCTC 2400
Qy 2923 TTCTCCGCTCTCAAGGTGACCAAGAAACCTGAAAGAGCTGGAACCTTAAGTGAACCTCGCTG 2982
Db 2401 TTCTCCGCTCTCAAGGTGACCAAGAAACCTGAAAGAGCTGGAACCTTAAGTGAACCTCGCTG 2460
Qy 2983 AGCCACTCTGAGTGAAGAGTCTTTGTAAGACCTTGAGACGCTCGCTGCTCTGAGAG 3042
Db 2461 AGCCACTCTGAGTGAAGAGTCTTTGTAAGACCTTGAGACGCTCGCTGCTCTGAGAG 2520
Qy 3043 ACCCTGCGGTGGCTGCTGCTGCTGCTCAAGCTGAGAGCTGAGAGACCTTGGCTTGGG 3102
Db 2521 ACCCTGCGGTGGCTGCTGCTGCTGCTCAAGCTGAGAGCTGAGAGACCTTGGCTTGGG 2580
Qy 3103 CTGAGAGCAACCAAGACCTGACCGAGCTGGAACCTGAGCTTCAATGCTCAACGAGTCT 3162
Db 2581 CTGAGAGCAACCAAGACCTGACCGAGCTGGAACCTGAGCTTCAATGCTCAACGAGTCT 2640
Qy 3163 GGAAGCAAAACCTTTGCGAGAGCTGAGACGCGGAGCTGCAAGCTCAAGGAGCTGAG 3222
Db 2641 GGAAGCAAAACCTTTGCGAGAGCTGAGACGCGGAGCTGCAAGCTCAAGGAGCTGAG 2700
Qy 3223 CTGCTCACTGAGTGGCTCAAGCTCTGATGCTGCTGCGAGACCTGCTGCTGATGCTC 3282
Db 2701 CTGCTCACTGAGTGGCTCAAGCTCTGATGCTGCTGCGAGACCTGCTGCTGATGCTC 2760
Qy 3283 AGCCCAACCTGAGAGAGTGAACCTGAGAGAGCAACCTGAGAGAGAGTGGAGCTGAG 3342
Db 2761 AGCCCAACCTGAGAGAGTGAACCTGAGAGAGCAACCTGAGAGAGAGTGGAGCTGAG 2820
Qy 3343 CTGCTCTGAGAGGCTCAAGGATCTGCTGCAAACTCATACGCTGAGGCTGAGACAG 3402
Db 2821 CTGCTCTGAGAGGCTCAAGGATCTGCTGCAAACTCATACGCTGAGGCTGAGACAG 2880
Qy 3403 ACAACTCTGAGTGAATGAGATGAGGAGAGAACTGAGAGGCTGAGAGAGAGAACTGAG 3462
Db 2881 ACAACTCTGAGTGAATGAGATGAGGAGAGAACTGAGAGGCTGAGAGAGAGAACTGAG 2940
Qy 3463 CTGCTCATCTTCAAGAGACGAGAAACCAAGTGTATGACCCCTCACTGAGAGGCTGAGATCG 3522

Db 2941 CTGCTCATCTTCAAGAGACGAGAAACCAAGTGTATGACCCCTACTGAGGCTGATACG 3000
Qy 3523 GAGAGATGAGATTAATGACATCTCTCACTCAAGCGGAGAGACTCGAGATCAAGAGGCG 3582
Db 3001 GAGAGATGAGATTAATGACATCTCTCACTCAAGCGGAGAGACTCGAGATCAAGAGGCG 3060
Qy 3583 GCTTCCCATGTTGCTCAGGCTTAATCTCAAACTCTGAGACGAGACAGATCTTCCCAATT 3642
Db 3061 GCTTCCCATGTTGCTCAGGCTTAATCTCAAACTCTGAGACGAGACAGATCTTCCCAATT 3120
Qy 3643 GCTGAGATTGAGAGAGAAAGCTCCCAAGAGTATGACCGGAGAACTTTGCGTGCCT 3702
Db 3121 GCTGAGATTGAGAGAGAAAGCTCCCAAGAGTATGACCGGAGAACTTTGCGTGCCT 3180
Qy 3703 TCTCTGCTCTTCAAGGAGAACCTGACATAGAGCTTTGGGAGCTGACGATGACTTGG 3762
Db 3181 TCTCTGCTCTTCAAGGAGAACCTGACATAGAGCTTTGGGAGCTGACGATGACTTGG 3240
Qy 3763 GGGCCCAAGGAGCTGAGCTAAGATGATGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3241 GGGCCCAAGGAGCTGAGCTAAGATGATGACAAAGAAAGAACTTGTACCGAGTT 3300
Qy 3823 CACTTCCCTGATGAGCTGCTCCTACCGCTGAGCCAAACGAGGCTCTGCTTGTATGAGA 3882
Db 3301 CACTTCCCTGATGAGCTGCTCCTACCGCTGAGCCAAACGAGGCTCTGCTTGTATGAGA 3360
Qy 3883 GAAAGGATGACCGTTGAGATTGAATCTGTGTGTGAGACAGATTCCTGGATGATCAAC 3942
Db 3361 GAAAGGATGACCGTTGAGATTGAATCTGTGTGTGAGACAGATTCCTGGATGATCAAC 3420
Qy 3943 CCAAGACACAGCTGATGATGATGAGGAGGCTCTGCTGAGACATCAAGGCTGAGCTGAGCT 4002
Db 3421 CCAAGACACAGCTGATGATGATGAGGAGGCTCTGCTGAGACATCAAGGCTGAGCTGAGCT 3480
Qy 4003 GTGAGAGTGTGACCTCCCTCACTTTGAGCTCTCAAGGAGGAGCAATGAGACATCC 4062
Db 3481 GTGAGAGTGTGACCTCCCTCACTTTGAGCTCTCAAGGAGGAGCAATGAGACATCC 3540
Qy 4063 CTGTTCCAAATGAGCCACTTTAAAGAGAGAGAGTCTCTGAGAAACCGACGAGGCTG 4122
Db 3541 CTGTTCCAAATGAGCCACTTTAAAGAGAGAGAGTCTCTGAGAAACCGACGAGGCTG 3600
Qy 4123 GAGCTGATCAATAGTCTTGAAGAAACCGAGCTTCTCCCTTGGAGAGTCTCTGAAA 4182
Db 3601 GAGCTGATCAATAGTCTTGAAGAAACCGAGCTTCTCCCTTGGAGAGTCTCTGAAA 3660
Qy 4183 ATGATCCATAATGAGCTGAGCTTCACTCCGCTCACTCTGAGTGTGCTTACACCGC 4242
Db 3661 ATGATCCATAATGAGCTGAGCTTCACTCCGCTCACTCTGAGTGTGCTTACACCGC 3720
Qy 4243 GTCCATCTTGAAGAGTCAACCTTCACTCTTCACTGATCCCAAGTGAATGCTCAATCGG 4302
Db 3721 GTCCATCTTGAAGAGTCAACCTTCACTCTTCACTGATCCCAAGTGAATGCTCAATCGG 3780
Qy 4303 AAGGAATCTGAGCTCTGCTATGAGAGCTTGAAGAAACCGAGCTTCTCGGAGTCTAC 4362
Db 3781 AAGGAATCTGAGCTCTGCTATGAGAGCTTGAAGAAACCGAGCTTCTCGGAGTCTAC 3840
Qy 4363 GTTGGCACTTGGAGATCAGGAGTCAAGGCTGCAATGAGAAAGCAAGAAAGATGAGACTG 4422
Db 3841 GTTGGCACTTGGAGATCAGGAGTCAAGGCTGCAATGAGAAAGCAAGAAAGATGAGACTG 3900
Qy 4423 GTGTGGAGGCTTGTGTAAGACAGAGATCTCATGCTGCACTACTCTGATCCCTCA 4482
Db 3901 GTGTGGAGGCTTGTGTGAAGACAGAGATCTCATGCTGCACTACTCTGATCCCTCA 3960
Qy 4483 GCGCGCATAGCCGTAATCTTCACTCTGAGATGCTCCCGAGTGTGCTGCTTGTGAGACAG 4542
Db 3961 GCGCGCATAGCCGTAATCTTCACTCTGAGATGCTCCCGAGTGTGCTGCTTGTGAGACAG 4020
Qy 4543 TATGAGAGCAGTATAGCCGAGTGAATCGTGTGAGGTTGTCTTGTGACAAACTGAT 4602
Db 4021 TATGAGAGCAGTATAGCCGAGTGAATCGTGTGAGGTTGTCTTGTGACAAACTGAT 4080

QY	301	GTCTTGAGAGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT	360
Db	331	GTCTTGAGAGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT	390
QY	361	CGGCTTATCTGTAAGTGGCTCTGGGATTTATAACTGGGTTCCGAATGCTGAATAGA	420
Db	391	CGGCTTATCTGTAAGTGGCTCTGGGATTTATAACTGGGTTCCGAATGCTGAATAGA	450
QY	421	GACGGTAAGAGCCAAAGCAAGACAGCACTGTTCTCTGGCTGGCTGATACCTCACAC	480
Db	451	GACGGTAAGAGCCAAAGCAAGACAGCACTGTTCTCTGGCTGGCTGATACCTCACAC	510
QY	481	CTGGGAACATCCCCAGACACCTCTTAACTCCGGGACAGATGGCTGGCGAGCTGG	540
Db	511	CTGGGAACATCCCCAGACACCTCTTAACTCCGGGACAGATGGCTGGCGAGCTGG	570
QY	541	GGCGGCTGGCTGTTACTTGGAGTTCTTGAAAGAGAGAGCTGAAGAGTTCCACTT	600
Db	571	GGCGGCTGGCTGTTACTTGGAGTTCTTGAAAGAGAGAGCTGAAGAGTTCCACTT	630
QY	601	CTGCTGSCAATAAAGGCACTCCAGAGCTCTTCGGGTGAGACACCGCTCAGCCAGAG	660
Db	631	CTGCTGSCAATAAAGGCACTCCAGAGCTCTTCGGGTGAGACACCGCTCAGCCAGAG	690
QY	661	AAGACAGTGGCATGAGTGGCTCTGTAAGTGGGTGAGTGGGAGCAGCGGCG	720
Db	691	AAGACAGTGGCATGAGTGGCTCTGTAAGTGGGTGAGTGGGAGCAGCGGCG	750
QY	721	TGGGACCTAGCCCTCCATCTTGGAGAGATGGGGCTGAGGTCACTGTGGCCCAAGCC	780
Db	751	TGGGACCTAGCCCTCCATCTTGGAGAGATGGGGCTGAGGTCACTGTGGCCCAAGCC	810
QY	781	CAGGAAGGGGCAAGGCACTCTCCCTCATTCCTCTACAGCCCAAGTGAACCCCACTG	840
Db	811	CAGGAAGGGGCAAGGCACTCTCCCTCATTCCTCTACAGCCCAAGTGAACCCCACTG	870
QY	841	TCTCCAGCCCAAGGCACTCTCCCTCATTCCTCTACAGCCCAAGTGAACCCCACTG	900
Db	871	TCTCCAGCCCAAGGCACTCTCCCTCATTCCTCTACAGCCCAAGTGAACCCCACTG	930
QY	901	GGGTGACCCCAAGGCTCAGAGAGAGAGGTTTGAAGACAGCTCTGACATCTGAGAGC	960
Db	931	GGGTGACCCCAAGGCTCAGAGAGAGAGGTTTGAAGACAGCTCTGACATCTGAGAGC	990
QY	961	CGCTGAGAGAAATCTGAGCTCACTCTCTACCAAGGCTCTTCCAGCTCCCGACAT	1020
Db	991	CGCTGAGAGAAATCTGAGCTCACTCTCTACCAAGGCTCTTCCAGCTCCCGACAT	1050
QY	1021	GAGTCTTCAAGCCAGAGTCAACCCAGCGCCCAATCCACAGAGTGTGGGAGCTGG	1080
Db	1051	GAGTCTTCAAGCCAGAGTCAACCCAGCGCCCAATCCACAGAGTGTGGGAGCTGG	1110
QY	1081	GGATCCCCACCTAGCCCAAGCTTGAACCCAGAGAGAGAGGCTCTGGGAGCCCAATGG	1140
Db	1111	GGATCCCCACCTAGCCCAAGCTTGAACCCAGAGAGAGAGGCTCTGGGAGCCCAATGG	1170
QY	1141	CTCTGATGATAAGCTAGAGATTTTACTACAGAAATCAGAGAAAGAGAGAGAGAA	1200
Db	1171	CTCTGATGATAAGCTAGAGATTTTACTACAGAAATCAGAGAAAGAGAGAGAGAA	1230
QY	1201	TCAGAGAAAGGAG	1260
Db	1231	TCAGAGAAAGGAG	1290
QY	1261	AGCTTACAGCCCAAGCCCAATGGAGAGCTTCTGTGAGAGAGAGCTTGTTCACA	1320
Db	1291	AGCTTACAGCCCAAGCCCAATGGAGAGCTTCTGTGAGAGAGAGCTTGTTCACA	1350
QY	1321	TGGCCCTGAGAAATGAGATTTTAAACCAAAATTCACACAGCTGCTACTTCTCAAGA	1380
Db	1351	TGGCCCTGAGAAATGAGATTTTAAACCAAAATTCACACAGCTGCTACTTCTCAAGA	1410
QY	1381	CTTCAACCCCAAGAGCCCAAGATCCCTGGCTCAAGAGAGCTGGCTGATTAATGAGAG	1440
Db	1411	CTTCAACCCCAAGAGCCCAAGATCCCTGGCTCAAGAGAGCTGGCTGATTAATGAGAG	1470
QY	1441	AATGAGAGATTTAATGAGATCAAGAGCTTTTGGGCCCAAGGCTGAGTATCCAGAA	1500
Db	1471	AATGAGAGATTTAATGAGATCAAGAGCTTTTGGGCCCAAGGCTGAGTATCCAGAA	1530
QY	1501	CTTGCATAGTCAATCTGACAGGAGGCTGAGATTTGGGAATGCAACATGAGCCAGAG	1560
Db	1531	CTTGCATAGTCAATCTGACAGGAGGCTGAGATTTGGGAATGCAACATGAGCCAGAG	1590
QY	1561	GTAAGAGAGCCCTGGGAGAGAGGCAAGCTGTATGGGAGCCGCTTCAGCATGTCTTAC	1620
Db	1591	GTAAGAGAGCCCTGGGAGAGAGGCAAGCTGTATGGGAGCCGCTTCAGCATGTCTTAC	1650
QY	1621	TTGAGCTGAGAGAGCTGGGAGGCTCCAGGTGTGAGTCTGCTGAGCTCATGGGAAA	1680
Db	1651	TTGAGCTGAGAGAGCTGGGAGGCTCCAGGTGTGAGTCTGCTGAGCTCATGGGAAA	1710
QY	1681	GATGGAGACCCCACTCCGCTCCCATTTAGACAGATCTGTCTAGGCCAGAGCGGCTCTC	1740
Db	1711	GATGGAGACCCCACTCCGCTCCCATTTAGACAGATCTGTCTAGGCCAGAGCGGCTCTC	1770
QY	1741	TTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1800
Db	1771	TTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1830
QY	1801	TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1860
Db	1831	TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1890
QY	1861	ACTATCTTCCAGAGGATCTTCTGATCAGGCTGAGCAAGCTCTGACAGAACTC	1920
Db	1891	ACTATCTTCCAGAGGATCTTCTGATCAGGCTGAGCAAGCTCTGACAGAACTC	1950
QY	1921	ATTCTCTTCTTGAAGAGGAGGCTTGGGTGAGAGGCTTGGGTGAGAGGCTTGGGTG	1980
Db	1951	ATTCTCTTCTTGAAGAGGAGGCTTGGGTGAGAGGCTTGGGTGAGAGGCTTGGGTG	2010
QY	1981	AAGGAATATTTCTACAGATTTTCAAGATGAAGAGCAATTAAGCTTAAAGGTTG	2040
Db	2011	AAGGAATATTTCTACAGATTTTCAAGATGAAGAGCAATTAAGCTTAAAGGTTG	2070
QY	2041	GTCAAATCAAAAGAGCTCTGGGCTCTGTCTTGTGGCTTGGGCTCTGGGCTGGCC	2100
Db	2071	GTCAAATCAAAAGAGCTCTGGGCTCTGTCTTGTGGCTTGGGCTCTGGGCTGGCC	2130
QY	2101	TGCACTTGGCTGATGAGAGATGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2160
Db	2131	TGCACTTGGCTGATGAGAGATGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2190
QY	2161	ACCAACACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGCCATTGGGACC	2220
Db	2191	ACCAACACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGCCATTGGGACC	2250
QY	2221	CAGCTCAGAGAGCTTGGCTCTGTGGCTGTGAGGAGCTTGGCAAAAAAGAGCCCTTTT	2280
Db	2251	CAGCTCAGAGAGCTTGGCTCTGTGGCTGTGAGGAGCTTGGCAAAAAAGAGCCCTTTT	2310
QY	2281	AGTCCAGATGACCTCAGAGAGATGGGTTGATGGGAGCATCTTCACTCTTTGAGAG	2340
Db	2311	AGTCCAGATGACCTCAGAGAGATGGGTTGATGGGAGCATCTTCACTCTTTGAGAG	2370
QY	2341	ATGGGATTTCTTAAG	2400
Db	2371	ATGGGATTTCTTAAG	2430
QY	2401	CAAGAGTTCTTTCAGAGATGCTATGCTTGAAGATGAGAGAGAGAGAGAGAGAT	2460
Db	2431	CAAGAGTTCTTTCAGAGATGCTATGCTTGAAGATGAGAGAGAGAGAGAGAGAT	2490
QY	2461	TCTAATTCATCATATGATTTGGAAAAAGAGCTGAGAGATGAGAAATACATGGCTGTT	2520

Db	2491	TCATATTGCAATCATATGATTTGGAAAAAGCGCTTAAGATATATGAAATACATATGCGCTGTTT	2550
Qy	2521	GGGGCATCAACCAACACGTTTCCCTAATTTGGGCTGTTAAGTATGAGGGGAGAGAGATG	2580
Db	2551	GGGGCATCAACCAACACGTTTCCCTAATTTGGGCTGTTAAGTATGAGGGGAGAGAGATG	2610
Qy	2581	GAGAAACATCTTTCACTGGCGGGCTGTTCAGGGGGAGAACCTGATGTCATGTGGTCCCGTCC	2640
Db	2611	GAGAAACATCTTTCACTGGCGGGCTGTTCAGGGGGAGAACCTGATGTCATGTGGTCCCGTCC	2670
Qy	2641	CTGCAAGCTGTGCTGCAGCCACACTCTCTGTGAGTCCCTCACTGCTGTGTAAGAACCTCG	2700
Db	2671	CTGCAAGCTGTGCTGCAGCCACACTCTCTGTGAGTCCCTCACTGCTGTGTAAGAACCTCG	2730
Qy	2701	AACAAAACGTTCTGTGACACAAAGTATGAGGCCCATTTGAAAGAAATGGGCATGTGTAGAA	2760
Db	2731	AACAAAACGTTCTGTGACACAAAGTATGAGGCCCATTTGAAAGAAATGGGCATGTGTAGAA	2790
Qy	2761	ACACACATGAGAGCTCTTAAGTGTGACATTTCTGCATTAATTCAAGCCGCACAGTGAAGAG	2820
Db	2791	ACACACATGAGAGCTCTTAAGTGTGACATTTCTGCATTAATTCAAGCCGCACAGTGAAGAG	2850
Qy	2821	CTTCAGCTGATTTGAGGGGACGAGCAGACAGATCAATGAGAGCCCAACATGTATGCTGT	2880
Db	2851	CTTCAGCTGATTTGAGGGGACGAGCAGACAGATCAATGAGAGCCCAACATGTATGCTGT	2910
Qy	2881	TTCAAGTGTGGTCCCAAGTCAACAGATGCTTAATTTGGCAGATTCTCTTCTCGTCTCAAGATC	2940
Db	2911	TTCAAGTGTGGTCCCAAGTCAACAGATGCTTAATTTGGCAGATTCTCTTCTCGTCTCAAGATC	2970
Qy	2941	ACCAAGAACTGTAAAGAGAGCTGAGCCTTAAGTGGAACTCTGCTGAGCACACTTGCAGTGAAG	3000
Db	2971	ACCAAGAACTGTAAAGAGAGCTGAGCCTTAAGTGGAACTCTGCTGAGCACACTTGCAGTGAAG	3030
Qy	3001	AGCTCTTTGTAAAGACCCCTGAGACGGCCCTGTGCGCTCCGAGAACCCCTGGGTTTGGCTGC	3060
Db	3031	AGCTCTTTGTAAAGACCCCTGAGACGGCCCTGTGCGCTCCGAGAACCCCTGGGTTTGGCTGC	3090
Qy	3061	TGTGGCCTCAAGCTGAGAGCTGACAGAGACCTTGTGCGCTGAGAGCCCAACAGACAC	3120
Db	3091	TGTGGCCTCAAGCTGAGAGCTGACAGAGACCTTGTGCGCTCCGAGAACCCCTGGGTTTGGCTGC	3150
Qy	3121	CTGAACCGAGCTGGAACCTGAAGCTTCAATGTGCTCAACGAGTGTGAGGCCAAACACTTTTGC	3180
Db	3151	CTGAACCGAGCTGGAACCTGAAGCTTCAATGTGCTCAACGAGTGTGAGGCCAAACACTTTTGC	3210
Qy	3181	CAGAGACTGANAACAGCCGAGCTGCAAGTACAGGACCTGCAAGCTGTGTACGTGTGGCTC	3240
Db	3211	CAGAGACTGANAACAGCCGAGCTGCAAGTACAGGACCTGCAAGCTGTGTACGTGTGGCTC	3270
Qy	3241	ACGTCTGAAGCTGTGCAGAGACCTGAGCCCTGTGTGTAGTGTGCAGGCCCAAGCTTAAGAG	3300
Db	3271	ACGTCTGAAGCTGTGCAGAGACCTGAGCCCTGTGTGTAGTGTGCAGGCCCAAGCTTAAGAG	3330
Qy	3301	CTAAGACTGACAGCAACCAACCTTGAGTACGTTTGGCGTGCAGCTCTTGTGAAGGGCTC	3360
Db	3331	CTAAGACTGACAGCAACCAACCTTGAGTACGTTTGGCGTGCAGCTCTTGTGAAGGGCTC	3390
Qy	3361	AGGCATCTGTGCTGTCAAACTCATACGCTTGGGGCTGGAACAGACACCTTGAATGATGAG	3420
Db	3391	AGGCATCTGTGCTGTCAAACTCATACGCTTGGGGCTGGAACAGACACCTTGAATGATGAG	3450
Qy	3421	ATGAGAGCAGAACTGAGGGGCTGTGAGAGAGAGAAACCTCAAGCTGTCTCATTTCAAGACGA	3480
Db	3451	ATGAGAGCAGAACTGAGGGGCTGTGAGAGAGAGAAACCTCAAGCTGTCTCATTTCAAGACGA	3510
Qy	3481	CGAAACCAAGTGTGATGACCCCTTACTGAGGGCTGTGATACGGGAGAGATGATATATGC	3540
Db	3511	CGAAACCAAGTGTGATGACCCCTTACTGAGGGCTGTGATACGGGAGAGATGATATATATGC	3570
Qy	3541	ACATCTCTACTCAAGCGGACAGACTCGGATTCAGAGAGGGGGGCTTCCCATGTTGCTCAG	3600

Db	3571	ACATCCCTCACTCAAGGGGAGAGAGCTGGATTCAGAGAGGGGGGCTTCCATGTTGCTCAG	3653
Qy	3601	GCTAATCTCAAACTCCTGAGCGTGAGCAGATCTTCCAAATGCTGATGATTGCA	3654
Db	3631	GCTAATCTCAAACTCCCTGAGCGTGAGCAGATCTTCCAAATGCTGATGATTGCGAGAG	3690
Qy	3655	-----GAGGAAGAGCTCCCGAGAGATGATACCGGTGGAATCTTGTACGGGCTTCTCCT	3708
Db	3691	AGCCACAGGAAGAGCTCCCGAGAGATGATACCGGTGGAATCTTGTACGGGCTTCTCCT	3750
Qy	3709	GCCCTCAAGGGGAGCTGCATACGAAGCTTTTGGGGAGCTGACGATGACTTCTGGGGCCCC	3768
Db	3751	GCCTCTAAGGGGAGCTGCATACGAAGCTTTTGGGGAGCTGACGATGACTTCTGGGGCCCC	3810
Qy	3769	ACGGGGGCTGTGGCTACGAGATGATTGACAAAGAAAGAACTTTTACCGATTCCTC	3828
Db	3811	ACGGGGGCTGTGGCTACGAGATGATTGACAAAGAAAGAACTTTTACCGATTCCTC	3870
Qy	3829	CCGTACTACGATCTTACCGGCTGCCCCAACCGGGTCTCGCTTTGATGAGAGAAAGCG	3888
Db	3871	CCGTACTACGATCTTACCGGCTGCCCCAACCGGGTCTCGCTTTGATGAGAGAAAGCG	3930
Qy	3889	GTGACCTGTAAGATTGAATCTGTGTGTGGAGCAAGTTCCTGGGTGATCAACCCACAG	3948
Db	3931	GTGACCTGTAAGATTGAATCTGTGTGTGGAGCAAGTTCCTGGGTGATCAACCCACAG	3990
Qy	3949	CACAGCTGAATGTGTGAGAGGCTCTGTGCTGAGACATCAAGGCTGAGCTGTGAGAA	4008
Db	3991	CACAGCTGAATGTGTGAGAGGCTCTGTGCTGAGACATCAAGGCTGAGCTGTGAGAA	4050
Qy	4009	GCTGTGACATCTCCCTCACTTTGTGTGCTCTCCAAAGGGGGCCATGTGACACATCCCTGTC	4068
Db	4051	GCTGTGACATCTCCCTCACTTTGTGTGCTCTCCAAAGGGGGCCATGTGACACATCCCTGTC	4110
Qy	4069	CAAAATGSCCACTTTAAAGAGAGGGGATGCTCCTGAGAAAGCAGCAGGAGTGAGACTG	4128
Db	4111	CAAAATGSCCACTTTAAAGAGAGGGGATGCTCCTGAGAAAGCAGCAGGAGTGAGACTG	4170
Qy	4129	CATCACATAGTTCTGAGAAACCCCAAGCTTCTCCCTCTGGAGAGTCTCTGAATATATC	4188
Db	4171	CATCACATAGTTCTGAGAAACCCCAAGCTTCTCCCTCTGGAGAGTCTCTGAATATATC	4230
Qy	4189	CATTAATGCCCTGGCTTCATTTCCGCTACCTCTGTGTGTGCTTACACACGGCTCAT	4248
Db	4231	CATTAATGCCCTGGCTTCATTTCCGCTACCTCTGTGTGTGCTTACACACGGCTCAT	4290
Qy	4249	CCGAGGAGAGTCACCTTCCACTCTACCTATCCAGTACCGTCCATTCG-----	4302
Db	4291	CCGAGGAGAGTCACCTTCCACTCTACCTATCCAGTACCGTCCATTCG-----	4350
Qy	4303	-----	4302
Db	4351	ATAGATATCTAGAAATGAATTCAGTTGTGGAAATCAACAAGCACCCTGAC	4410
Qy	4303	-----	4302
Db	4411	CCAATTATATGAGGCTGCTGTTACACTGTGTCTGGGTCTGGTTCAAGGATGCTGAATA	4470
Qy	4303	-----AAGAACTGAGAGCTCTGCTATACGAAGCCCTGGAAGAGCAAGCTTCTGGAG	4356
Db	4471	CTCCCAAGGAAGCTGAGAGCTCTGCTATACGAAGCCCTGGAAGAGCAAGCTTCTGGAG	4530
Qy	4357	TTCTACCTTGGCCACTTGGGATCAGGAGTACAGGCTGCACTGTAAGACAAAGAAAGATGAG	4416
Db	4531	TTCTACCTTGGCCACTTGGGATCAGGAGTACAGGCTGCACTGTAAGACAAAGAAAGATGAG	4590
Qy	4417	ACTTGGTGTGGAGAGGCTTGTGTGAACCAAGGA 4449	
Db	4591	ACTTGGTGTGGAGAGGCTTGTGTGAACCAAGGA 4623	

RESULT 8
AX089761

LOCUS AX089761 4422 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 1 from Patent WO0116170.
 ACCESSION AX089761
 VERSION AX089761.1 GI:13443933
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 Reed, J.C.
 Card proteins involved in cell death regulation
 Patent: WO 0116170-A 1 08-MAR-2001;
 The Burnham Institute (US)
 JOURNAL Location/Qualifiers
 FEATURES
 Source
 1. 4422
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1. 4422
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC34899.1"
 /db_xref="GI:13443934"
 CDS
 /translation="MAGGAWRLACVLEPLKKEELKEFOILLANKHSSSSGPTPAO
 PERTSGMEVASVLAOYGBORANDLAHTMEQWELSLCAOABGAGHSSSPYSBE
 PHIGSPQPTSTAVLMFMIHELPACTOGSEBRVRLDPTSGRMEISASLYOL
 PSSPDHSPQESNPASTVAVLGSWSPQPSLAPEAGPGLTOWELSGIYYE
 IRRERKESKESNPAAVAVGTPOASTLSQPHHPEPVSRLSTCWPKWEDFNO
 KFTOLLQRPHPVSODPLVRSWPDYENRGHLIIRDLFGDLTOPRIYLQD
 AAGIKETLRQVSEAWRGOLYGDROHYFYSRLAOSKVSYLELIGDGTARP
 APIROLISPERLLFLIDYDEPQWTLQESSELCLMSQPOPADALGSLKTIIP
 EAPFLITARTTALONLIPSLERARWVLEPSSSRKRYTRFTIDRQAIRARLVK
 SNELWMLCVPMWSMLACTLMQOMRKLTLYSTTTTLCLATYLAOLQOPLD
 QLEDLGSLAEGIMQKTLFSPDRLRGHGDGAIISTFLMGILQEHPIPLSYSLIL
 CPOEFPAASGVLEDEKRGKHSNCIIDLKTLAAYIGLFGASITRFLGLSDG
 ERMENEFHCRLSGRMLQMWPSLOLLPHLSLESLHCLYERNTKFLQWAAHPEE
 MCMCVETMDLICTPCKFSRHYKQLILEGRHRETPPTWVLFRTVYDAYVO
 ILSVLTAVTNLKELDLSGNSLSHAYKSLCKTLRRRCLETLRLGGLTABDCD
 LAFLGRANQTLDELDSFNVLTDAAGHLCQRLRQPSCKLQRLQVSCGLQSCCDL
 ASVLASPSLKELDLQNNLDDVGRLLCGLRHPACKLRLGLDQTLTDEKOEEL
 ALDOEKPOLILFSRKSQVMTPEGLDTGMSNSTSLKQRLGSEBAASHVQANLK
 LLDVSKFLPIAIESSPEVIVPELLCVSPASQDHLTKRPLGIDDDFMGPRPVAT
 EVDKRNLYRHPVPAVSRYRMYTGCCTPMBEAVTYEIRFCWDOGLGINSQMSM
 VAEPLDILKAPGAVEAVHLPHFVALQGGVDTSLFWAFKESGMLLEPARVELH
 IVLENPFSPLGVLLKMIHNALEPIPTSVLLYHRVPEBVEFHLILPSDQIRKA
 IDLEMFQEVRIHKPPPLPLVWGCRTVSGSGMLETILPELEICTRSPGDDQF
 SEFVGHLSGIRLOVKDKDETLVMEALVKPGLDPAATLILPARLAVSPDLAPOL
 LHFVDOYREOLIAVTSVEVYLDLHGQVLSOEYERVLLENTRPSQMKLFLSLQSW
 DRCKDQYQALKEITHPLIMELMEKSKKGLPLFS"

BASE COUNT 1043 a 1224 c 1240 g 915 t
 ORIGIN
 Query Match 76.2%; Score 4148; DB 6; Length 4422;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

QY 763 TCACTGNGGCGCCAAAGGCCAGAAAGGGGCAAGGCCACTCTCCATTCATCCCTACAGCCCA 822
 DB 241 TCACTGTGCGCCCAAGGCCAGAAAGGGGCAAGGCCACTCTCCATTCATCCCTACAGCCCA 300
 QY 823 AGTGAACCCCACTGGGGTCTCCAGCAACCCACTCCACCTGCAAGTCTAATGCTGGG*882
 DB 301 AGTGAACCCCACTGGGGTCTCCAGCAACCCACTCCACCTGCAAGTCTAATGCTGGG 360
 QY 883 ATCCATGAATTGCGCGGGGGTGCACCCAGGGCTCAGAGAAAGGTTTGAACAGCTG 942
 DB 361 ATCCATGAATTGCGCGGGGGTGCACCCAGGGCTCAGAGAAAGGTTTGAACAGCTG 420
 QY 943 CCGACACATCTGAGCGCGCTGAGAGAAATCTGCGCTCACTCCCTTACCAAGCTCT 1002
 DB 421 CCGACACATCTGAGCGCGCTGAGAGAAATCTGCGCTCACTCCCTTACCAAGCTCT 480
 QY 1003 CCAAGCTCCCAAGCACTAGTCTCCAGAGCAGAGTCAACCAAGCCGCCACATCCACA 1062
 DB 481 CCAAGCTCCCAAGCACTAGTCTCCAGAGCAGAGTCAACCAAGCCGCCACATCCACA 540
 QY 1063 GCAAGTCTGGGAGCTGGGGATCCCACTCAGCCCAAGCTTACGACCCAGAGCAGAG 1122
 DB 541 GCAAGTCTGGGAGCTGGGGATCCCACTCAGCCCAAGCTTACGACCCAGAGCAGAG 600
 QY 1123 GCTCTGGGAGCCCAATGGCCTCTGATGAACGTCAAGGAATTTACTACAGAAATCAGA 1182
 DB 601 GCTCTGGGAGCCCAATGGCCTCTGATGAACGTCAAGGAATTTACTACAGAAATCAGA 660
 QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATAGGCGAGGTTGATGAGACG 1242
 DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATAGGCGAGGTTGATGAGACG 720
 QY 1243 CCCCACAGGCGCACACAGCTTACAGCCCAACCAACCAATGGAGCTTCTGTAGA 1302
 DB 721 CCCCACAGGCGCACACAGCTTACAGCCCAACCAACCAATGGAGCTTCTGTAGA 780
 QY 1303 GAGAGCCTGTGTTCCATGAGGCGCTGGAATAAGAGATTTTAAACCAAAATTCACAG 1362
 DB 781 GAGAGCCTGTGTTCCATGAGGCGCTGGAATAAGAGATTTTAAACCAAAATTCACAG 840
 QY 1363 CTGCTACTTCTAAGAAAGCTCAACCCAGAGCAAGATCCCTGTCAAGAGAGCTGG 1422
 DB 841 CTGCTACTTCTAAGAAAGCTCAACCCAGAGCAAGATCCCTGTGTCAAGAGAGCTGG 900
 QY 1423 CTTGATTATGTGAGAGAAATCAGAGCAATTTAATTGATTCAGAGACTTATTTGGCCCA 1482
 DB 901 CTTGATTATGTGAGAGAAATCAGAGCAATTTAATTGATTCAGAGACTTATTTGGCCCA 960
 QY 1483 GGCCTGATACCCAAAGAACTCGCATATGCTACTGAGGGGGCTGCTGAAATTTGGGAG 1542
 DB 961 GGCCTGATACCCAAAGAACTCGCATATGCTACTGAGGGGGCTGCTGAAATTTGGGAG 1020
 QY 1543 TCAACTGTGCGAGGAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGAGCCG 1602
 DB 1021 TCAACTGTGCGAGGAGTGAAGAAAGCTGGGGGAGAGCCAGCTGTATGGGAGCCG 1080
 QY 1603 TTCCAGCATGTCTTCTAATCTTCACTGAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1662
 DB 1081 TTCCAGCATGTCTTCTAATCTTCACTGAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1140
 QY 1663 GCTGAGCTCATGGGAAAGATGGGACAGGCCATCCGGCTCCATTAGACAGATCTGTCT 1722
 DB 1141 GCTGAGCTCATGGGAAAGATGGGACAGGCCATCCGGCTCCATTAGACAGATCTGTCT 1200
 QY 1723 AGGCGAGAGCGGCTGCTTCACTCTGATGTGTAGATGAGCAAGATGGGCTTTGACAG 1782
 DB 1201 AGGCGAGAGCGGCTGCTTCACTCTGATGTGTAGATGAGCAAGATGGGCTTTGACAG 1260
 QY 1783 GAGCCGAGTCTGAGCTCTGTCTGCACTGAGAGCCAGCAAGCCGGGAGATGCACTGCTG 1842
 DB 1261 GAGCCGAGTCTGAGCTCTGTCTGCACTGAGAGCCAGCAAGCCGGGAGATGCACTGCTG 1320
 QY 1843 GCGAGTTGTGCGGAAAACTATTACTTCCGAGGCACTCTCTGATCAAGGCTCGAGCC 1902

Db 1321 GGCAAGTTTGTGGGAAAACCTAATCTTCCGAGGCAACCTTCTGATCAAGGCTCGAAC 1380
 Qy 1903 AACAGCTCTGAGAAACCTATTCCTTCTTTGAGACAGGACGTTGGGTAGAGTCTCGGAG 1962
 Db 1381 AAGAGCTGTGAGAAACCTATTCCTTCTTTGAGACAGGACGTTGGGTAGAGTCTCGGAG 1440
 Qy 1963 TTCTCTGAGTCCAGCAGAAAGAAATTTTCTACATATTTTCAAGATGAAAGGACGCA 2022
 Db 1441 TTCTCTGAGTCCAGCAGAAAGAAATTTTCTACATATTTTCAAGATGAAAGGACGCA 1500
 Qy 2023 ATTAGAGCCTTTAGTGTGCTCAATCAACAAAGAGCTCGGGCCGTGTGCTTGAGCC 2082
 Db 1501 ATTAGAGCCTTTAGTGTGCTCAATCAACAAAGAGCTCGGGCCGTGTGCTTGAGCC 1560
 Qy 2083 TGGGTCTCTGGCTGGCTGACCTTGTGCTGATGACAGATGAAAGGCAAGAAACTC 2142
 Db 1561 TGGGTCTCTGGCTGGCTGACCTTGTGCTGATGACAGATGAAAGGCAAGAAACTC 1620
 Qy 2143 ACACCTACCTTCCAGACACACACACACCTTGTCTACATTAACCTTCCAGGCTCTCCAA 2202
 Db 1621 ACACCTACCTTCCAGACACACACACACCTTGTCTACATTAACCTTCCAGGCTCTCCAA 1680
 Qy 2203 GCTCAGCCATTTGGGAGCCCAAGCTCAGAGACCTGCTGTGGCTGTGAGGAGCATCTGG 2262
 Db 1681 GCTCAGCCATTTGGGAGCCCAAGCTCAGAGACCTGCTGTGGCTGTGAGGAGCATCTGG 1740
 Qy 2263 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGTTAGTGGGACATC 2322
 Db 1741 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGTTAGTGGGACATC 1800
 Qy 2323 ATCTCAGCTTCTTGAAGATGGTATTTCTTGAAGACACCCCATCTCTGAGCTACAGC 2382
 Db 1801 ATCTCAGCTTCTTGAAGATGGTATTTCTTGAAGACACCCCATCTCTGAGCTACAGC 1860
 Qy 2383 TTCAATCACCTCTGTTTCCAAAGATCTTGTGACAGATGCTGATGCTTGAAGATGAG 2442
 Db 1861 TTCAATCACCTCTGTTTCCAAAGATCTTGTGACAGATGCTGATGCTTGAAGATGAG 1920
 Qy 2443 AAGGGAGAGTAAACATTTCTAATTCATCATATGATTTGAAAAAGACCTTGAAGCATAT 2502
 Db 1921 AAGGGAGAGTAAACATTTCTAATTCATCATATGATTTGAAAAAGACCTTGAAGCATAT 1980
 Qy 2503 GGAATCATGAGCTGTTTGGGGGATCAACACACCTTCTATTTGGGCTCTTAAATGAT 2562
 Db 1981 GGAATCATGAGCTGTTTGGGGGATCAACACACCTTCTATTTGGGCTCTTAAATGAT 2040
 Qy 2563 GAGGGGAGAGAGATGAGAAACATCTTCACTGCGGCTGTCTCAGGGGAGAACTG 2622
 Db 2041 GAGGGGAGAGAGATGAGAAACATCTTCACTGCGGCTGTCTCAGGGGAGAACTG 2100
 Qy 2623 ATGCAATGAGTCCGCTCCCTGAGCTGTGCTGACAGACACCTCTGAGTCCCTCAC 2682
 Db 2101 ATGCAATGAGTCCGCTCCCTGAGCTGTGCTGACAGACACCTCTGAGTCCCTCAC 2160
 Qy 2683 TGGTTTGAAGACTGAGAAACAAAGCTTCTGACACAAATGATGGCCCATTTGAGAA 2742
 Db 2161 TGGTTTGAAGACTGAGAAACAAAGCTTCTGACACAAATGATGGCCCATTTGAGAA 2220
 Qy 2743 ATGGGATGATGTAGAAACAGACATGAGAGCTCTTATGTGCACTTTTGCATTTAAATTC 2802
 Db 2221 ATGGGATGATGTAGAAACAGACATGAGAGCTCTTATGTGCACTTTTGCATTTAAATTC 2280
 Qy 2803 AGCGGCAAGTGAAGAGCTTCACTGATTTGAGGGCAGGACAGACATCAATGAGAGC 2862
 Db 2281 AGCGGCAAGTGAAGAGCTTCACTGATTTGAGGGCAGGACAGACATCAATGAGAGC 2340
 Qy 2863 CCAACATGATTAATCTGTTCAAGTGGTCCAGTCAAGATGCTTAAATGAGATTC 2922
 Db 2341 CCAACATGATTAATCTGTTCAAGTGGTCCAGTCAAGATGCTTAAATGAGATTC 2400
 Qy 2923 TTCTCGTCTCAAGTCAACAGAACTGAAGAGCTGAGCTTAAATGAGATTCGCTG 2982

Db 2401 TTCTCGTCTCAAGTCAACAGAAACCTGAAGAGCTGAGACCTTAAGTGAACCTGCTG 2460
 Qy 2983 AGCCACTGTGAGTAAGAGTCTTTGTAAAGACCTTGAAGACGCTCTGCTGCTCTGAG 3042
 Db 2461 AGCCACTGTGAGTAAGAGTCTTTGTAAAGACCTTGAAGACGCTCTGCTGCTCTGAG 2520
 Qy 3043 ACCCTGCGGTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 3102
 Db 2521 ACCCTGCGGTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2580
 Qy 3103 CTGAGAGCCACAGACCTGACCTGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 3162
 Db 2581 CTGAGAGCCACAGACCTGACCTGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 2640
 Qy 3163 GAGCCAAACACCTTTGCCAGAGCTGAGACAGCCGAGCTGAGACCTTCAAGGACTGAG 3222
 Db 2641 GAGCCAAACACCTTTGCCAGAGCTGAGACAGCCGAGCTGAGACCTTCAAGGACTGAG 2700
 Qy 3223 CTGCTCAGTGTGGCTCAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 3282
 Db 2701 CTGCTCAGTGTGGCTCAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2760
 Qy 3283 AGCCCAAGCTGAGAGCTGAGACCTGAGCAGAACCACTGATGACCTTGGAGTGA 3342
 Db 2761 AGCCCAAGCTGAGAGCTGAGACCTGAGCAGAACCACTGATGACCTTGGAGTGA 2820
 Qy 3343 CTGCTCTGTGAGGAGCTCAGGATCTGTGCTGCAAACTCATACGCTGTGGGCTGAGCAG 3402
 Db 2821 CTGCTCTGTGAGGAGCTCAGGATCTGTGCTGCAAACTCATACGCTGTGGGCTGAGCAG 2880
 Qy 3403 ACACCTGAGTGAATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3462
 Db 2881 ACACCTGAGTGAATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
 Qy 3463 CTGCTCATCTTCAAGAGACGAGAAACCAAGTGTGATGACCTCTACTGAGGAGCTGTGATAG 3522
 Db 2941 CTGCTCATCTTCAAGAGACGAGAAACCAAGTGTGATGACCTCTACTGAGGAGCTGTGATAG 3000
 Qy 3523 GAGAGATGATGAATGAGACATCTCTCATCAAGCGGAGAGACCTGAGATCAGAGAGGCG 3582
 Db 3001 GAGAGATGATGAATGAGACATCTCTCATCAAGCGGAGAGACCTGAGATCAGAGAGGCG 3060
 Qy 3583 GCTTCCAGTGTGCTCAGGCTAATCTCAACCTCTGAGAGCTGAGAGATCTTCCCAAT 3642
 Db 3061 GCTTCCAGTGTGCTCAGGCTAATCTCAACCTCTGAGAGCTGAGAGATCTTCCCAAT 3120
 Qy 3643 GCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACTCTTGTGCTGCT 3702
 Db 3121 GCTGAGATTGACAGAGAAAGCTCCCAAGAGTATGACCGGTGAACTCTTGTGCTGCT 3180
 Qy 3703 TCTCTGCTCTCAAGGGGACCTGCAATGAAAGCTTTTGGGAGCTGACGATGACTTCTGG 3762
 Db 3181 TCTCTGCTCTCAAGGGGACCTGCAATGAAAGCTTTTGGGAGCTGACGATGACTTCTGG 3240
 Qy 3763 GGGCCCAAGGGGCTGTGCTGCTGAGAGTGAACAAAGAAAGAACTTGAACGAGT 3822
 Db 3241 GGGCCCAAGGGGCTGTGCTGCTGAGAGTGAACAAAGAAAGAACTTGAACGAGT 3300
 Qy 3823 CACTTCCCTGTAGCTGCTCTCAACGCTGAGCCCAACAGCGGTCTGTGCTTGTGATGAG 3882
 Db 3301 CACTTCCCTGTAGCTGCTCTCAACGCTGAGCCCAACAGCGGTCTGTGCTTGTGATGAG 3360
 Qy 3883 GAAAGGATGACCTGTTGATGAATTTCTGTGTGTGAGACCAATCTCTGGGTGATCAAC 3942
 Db 3361 GAAAGGATGACCTGTTGATGAATTTCTGTGTGTGAGACCAATCTCTGGGTGATCAAC 3420
 Qy 3943 CCAAGACAGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4002
 Db 3421 CCAAGACAGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3480
 Qy 4003 GTGGAAGCTGTGACCTCTCTCACTTGTGCTCTTCAAGGGGAGGAGGAGGAGGAGGAGGAG 4062
 Db 3481 GTGGAAGCTGTGACCTCTCTCACTTGTGCTCTTCAAGGGGAGGAGGAGGAGGAGGAGGAG 3540

QY	4063	CTGTTCCAAATAGGCGCACTTTAAAGAGAGAGGAGATGCTCTGGAGAGACCGCAGGGTG	4122
Db	3541	CTGTTCCAAATAGGCGCACTTTAAAGAGAGAGAGATGCTCTGGAGAGACCGCAGGGTG	3600
QY	4123	GAGCTGCATCATATAGTCTTGGAAAAACCCAGCTTCTCCCTTGGAGAGTCTCTGAAA	4182
Db	3601	GAGCTGCATCATATAGTCTTGGAAAAACCCAGCTTCTCCCTTGGAGAGTCTCTGAAA	3660
QY	4183	ATGATCCATATAGCTCCCTGCGCTTCAATTCGCCGTACCTCTGTGTGTGTGCTTTACACCGC	4242
Db	3661	ATGATCCATATAGCTCCCTGCGCTTCAATTCGCCGTACCTCTGTGTGTGTGCTTTACACCGC	3720
QY	4243	GTCCATCTCTGAGGAAGTACCTTCCACCTCAACCTGATCCCAAGTGACTGTCCATTGCG	4302
Db	3721	GTCCATCTCTGAGGAAGTACCTTCCACCTCAACCTGATCCCAAGTGACTGTCCATTGCG	3780
QY	4303	-----	4302
Db	3781	AAGGCCATATAGTATCTAGAAAATGAATAATCCAGTTTGTGGAAATCCACAGCCACCCCG	3840
QY	4303	-----	4302
Db	3841	CTGACCCCACTTTATATAGGCGTGTGTGTTACATGTGTCTGGGTGTGTTCAAGGATGCTG	3900
QY	4303	-----	4302
Db	3901	GAAATACTCCCGCAAGAACTGAGACTCTGTATGAAAGCCCTGGAGAAAGACACACTGTT	3960
QY	4351	TCGGAGTTCTACGTTGGCCACTTGGAGATCAGAGATCAGGCTGCAAGTGAAGACAGAAA	4410
Db	3961	TCGGAGTTCTACGTTGGCCACTTGGAGATCAGAGATCAGGCTGCAAGTGAAGACAGAAA	4020
QY	4411	GATGAGACTCTGTGTGTGGAGAGCCTTGTGTGAACACAGAGATCTCATGTCTGCACTACT	4470
Db	4021	GATGAGACTCTGTGTGTGGAGAGCCTTGTGTGAACACAGAGATCTCATGTCTGCACTACT	4080
QY	4471	CTGATCTCCCTCAGGCGCGCATATAGCCGTACCTTCAGATATGCCCGCAGATTGCTGAC	4530
Db	4081	CTGATCTCCCTCAGGCGCGCATATAGCCGTACCTTCAGATATGCCCGCAGATTGCTGAC	4140
QY	4531	TTTGTGAGCAAGTATCGAAGACAGCTGATAGCCCGAAGTACATCATGTGTGAGAGTTGTCTTG	4590
Db	4141	TTTGTGAGCAAGTATCGAAGACAGCTGATAGCCCGAAGTACATCATGTGTGAGAGTTGTCTTG	4200
QY	4591	GACCAACTGCATGAGACAGGTGCTGAGCCAGAGACAGTACGAGAGGGTGTCTGAGAAC	4650
Db	4201	GACCAACTGCATGAGACAGGTGCTGAGCCAGAGACAGTACGAGAGGGTGTCTGAGAAC	4260
QY	4651	ACGAGGCGCCAGCCAGATGCGGAACTGTTCAGCTTGAACCAATCTTGGACCGGAAATGTC	4710
Db	4261	ACGAGGCGCCAGCCAGATGCGGAACTGTTCAGCTTGAACCAATCTTGGACCGGAAATGTC	4320
QY	4711	AAAGATGAGCTCTTACCAAGCCCTGAGAGAGACCATCTCACTCATTTATGAACTCTGG	4770
Db	4321	AAAGATGAGCTCTTACCAAGCCCTGAGAGAGACCATCTCACTCATTTATGAACTCTGG	4380
QY	4771	GAGAGAGGCGAGCAAAAAGGACTCTCTGCACTCAGCAGCTTA	4812
Db	4381	GAGAGAGGCGAGCAAAAAGGACTCTCTGCACTCAGCAGCTTA	4422
RESULT 9			
AF229060			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
4422 bp mRNA linear PRI 02-FEB-2001 Homo sapiens NAC-beta splice variant (NAC) mRNA, complete cds, alternatively spliced. AF229060 AF229060.1 GI:12656106			

REFERENCE	FEATURES
1 (bases 1 to 4422)	Location/Qualifiers
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1..4422
Chu, Z.-L., Pio P., Xie, Z., Godzik, A. and Reed, J. C.	/organism="Homo sapiens"
NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c pathway for apoptosis	/mol_type="mRNA"
Unpublished	/db_xref="taxon:9606"
2 (bases 1 to 4422)	/chromosome="11"
Chu, Z.-L. and Reed, J. C.	/cell_type="activated Jurkat T cells"
Direct Submission	1..4422
Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA	/gene="NAC"
	1..4422
	/gene="NAC"
	/note="Apaf-1/Ced-4 family apoptosis regulator; contains NB-ARC domain, CARD domain, and leucine-rich repeats (LRR)"
	/codon_start=1
	/product="NAC-beta splice variant"
	/protein_id="AAK00749.1"
	/db_xref="GI:12656107"
	/translation="MAGGAGRLACTYLFKKELKEFOLLANKHSSSGETPAQ PEKTSMEVASVIVAYGEORAWDLAHTMEQGLSLCAQAGEGASHSPSPSE PHGSSQEPSTAVIMPIWHELPAQCTGSEERVLRLPDSQRWRLEASLIYQAL PSSDPESE6SESPNAPSTAVLGSWGSPPQELARDECAETQWPLETSGIYYE IRREBEKSEKGRPPMAAVGTPROAHTSLQPHHMEPEVRSLSCTPMKAEDEFO KPTOLLRLRPHRSODPLTKYRSMPEVYENRRHLLEIRLPEGLDTPSRVILVIG AAIGISTLARQYKEMWGRGLYGDPRFQHYFTRSCHELAOSKYSALAEIKGKSTP APRQLSLRPERLLFTLDGDEPQWGLQVSESLSCAMSDPOADLALGSLLEKTTIP EASFLIARTATLALQNLPSLEQARWEVGLFSSSSKREYFYRTFERQALIRAFVAK SNEELWALCVPMVSWLACTCLMQMKRREKLTITKTTTLLCLHYALQALQOPGLP QLEDLCSLAEGIMCKTKLTFSPDDLKRGIDGAIISTFGLTGILQSHPIPLSSFTIL CFDEPFAAMSVYLEDKRGKSHNCIIDEKTELAEGIHGLFASSTRFLGLSDEG ERMENINFCRLRSQGRINQMWPSLQLLIQHSLSESLHCLYEIRNTKFLQWNAHEP MNCVETDEMLVCTPCITKSRMYKQLQILBGRQHSSTSPMYVLFRRVPTVADWQ ILTSVKVTKNLEKLDLSDGNSLSHSAVKSICKLRPRCLLETYRLAAGCGLTAECDL LAGLGAAPQTLLELDLSPFLVTLTDGAKHLQGLRQPSCKLQRLQVSCGLTSCODL ASTLASPSLKELDLQNNMLDVGAVLLCEGLHPACKLRILGLDITSLDEMRQELR ALAEORCOLLIIFSRKPSWTFTEGDLTGMSNSTSLKRORGSEBAASHVAQANIK LIDVSTIFPIAELIAEESSEPVVVELLCVSPASQGDILHKPLGTDPMFGTGPVAT EYVDKKNLIRYHFPVAGSTRMPTGLCYMEAAVYVIEFCWDOFLSEINQHSMM VAPDLILKAEPGAVEAHLPHFVALQGGHVALLSOMAFKEEKMILKPAVELHL VILENVSFSPGLVALKMINHARFIPVTSVLLYHVAHEVFPHLILPDSQIRVA IDILEMKFCQVRIRHPPPLTFLYMGCRVYVSGSGSMLELIPLELICYSPBEDOLF SEYVYRLLGSGIRLQVYDKKDELTMEVALYKSPDMLPATLILPPAIIYVSPIDAPOL LHEDVYQREGLIARVTSVEVLDKLHGQVLSQGYEYRVALENTRPSOMRKLFLSQSM DRCKDGYQALKEETHPHLMELMEKSGKKGLPLSS"
BASE COUNT	1043 a 1224 c 1240 g 915 t
ORIGIN	
Query Match	76.2%; Score 4148; DB 9; Length 4422;
Best Local Similarity	97.0%; Pred. No. 0;
Matches 4290; Conservative	0; Mismatches 0; Indels 132; Gaps 1;
Y	523 ATGGCTGGCGGAGCTTGGGGCGCGCTTGCTGTATCTTGGAGTTCTCGAAGAAAGAGAG 582
Db	1 ATGGCTGGCGGAGCTTGGGGCGCGCTTGCTGTATCTTGGAGTTCTCGAAGAAAGAGAG 60
Y	583 CTGAAGAGATTTCAGCTTTCGTGCTGCCTAAAGAGCACTCCAGAGAGCTTTGGGGGTAG 642
Db	61 CTGAAGAGATTTCAGCTTTCGTGCTGCCTAAAGAGCACTCCAGAGAGCTTTGGGGGTAG 120
Y	643 ACAACCGCTCAGCCAGAAAGACGAGTGGCATGAGAGTGGCTCTGTAACCTGTGTGCTAG 702

Db	121	ACACCCGCTCAGCCAGAGAAACAGAGTGGATGAGAGTGGGCTCTGTAACCTGGTGGCTCAG	180
Qy	703	TATGGGAGAGACCGGGGCTCTGGGACCTTAGCCCTTCATACCTGGAGACAGATGGGCTGAG	762
Db	181	TATGGGGGAGACGGGGCTGGGACCTTAGCCCTTCATACCTGGAGACAGATGGGCTGAG	240
Qy	763	TCATCTGTGGCCCAAGCCCAAGAAAGGGGACGGCCATCTCCCTCATTTCCCTCAACGCCA	822
Db	241	TCATCTGTGGCCCAAGCCCAAGAAAGGGGACGGCCATCTCCCTCATTTCCCTCAACGCCA	300
Qy	823	AGTGAACCCCACTGGGGTCTCCAGCCAAACCCATCCACCGGACAGTCTAATGGCCCTGG	882
Db	301	AGTGAACCCCACTGGGGTCTTCACAGCCAAACCCATCCACCGGACAGTCTAATGGCCCTGG	360
Qy	883	ATCCATGAATTTGCCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGACAGCTG	942
Db	361	ATCCATGAATTTGCCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGACAGCTG	420
Qy	943	CCTGACACATCTGGACCGCCGCTGGAGAGAAATCTTGCTCTCATCTCTTACCAAGCTCTT	1002
Db	421	CCTGACACATCTGGACCGCCGCTGGAGAGAAATCTTGCTCTCATCTCTTACCAAGCTCTT	480
Qy	1003	CCAAGCTCCCGAGCCATGAGTCTCCAGCCAGAGGTACCCACAGGCCCCCAATCCACA	1062
Db	481	CCAAGCTCCCGAGCCATGAGTCTTCAGCCAGAGGTACCCACAGGCCCCCAATCCACA	540
Qy	1063	GCAGTGTGGGGAAGCTGGGAGATCCCACTCAGCCCACTTAGCAACCCAGAGAGCAGAG	1122
Db	541	GCAGTGTGGGGAAGTGGGAGATCCCACTCAGCCCACTTAGCAACCCAGAGAGCAGAG	600
Qy	1123	GCTCCCTGGGACCCCAATGGCTCTTGAGTAAACGTGAGAAATTTACTACAGAAATCAGA	1182
Db	601	GCTCCCTGGGACCCCAATGGCTCTTGAGTAAACGTGAGAAATTTACTACAGAAATCAGA	660
Qy	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCAAGGCCCCCATGCGGACGGTGGTGGAAAG	1242
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGCAAGGCCCCCATGCGGACGGTGGTGGAAAG	720
Qy	1243	CCCCCAACGGGGCACACAGCCTCAGCGCCCAACCAACCCATGGGAGCCTTCTGTGAGA	1302
Db	721	CCCCCAACGGGGCACACAGCCTCAGCGCCCAACCAACCCATGGGAGCCTTCTGTGAGA	780
Qy	1303	GAGAGCCTCTGTTCACATGGCCTCTGAAAAATGAGATTTTAAACMAAAATTCACAG	1362
Db	781	GAGAGCCTCTGTTCACATGGCCTCTGAAAAATGAGATTTTAAACMAAAATTCACAG	840
Qy	1363	CTGCTACTCTTACAAAGACTCACCACCAAGACCAAGATCCCTGGTCAAGAGAGCTGG	1422
Db	841	CTGCTACTCTTACAAAGACTCACCACCAAGACCAAGATCCCTGGTCAAGAGAGCTGG	900
Qy	1423	CCTGATTTATGGAGAGAAATCAGAGCAATTTAATGAGATCAGAGCTTATTTGGGCCA	1482
Db	901	CCTGATTTATGGAGAGAAATCAGAGCAATTTAATGAGATCAGAGCTTATTTGGGCCA	960
Qy	1483	GGCCTGATACCAGAAACCTCGCATAGTCACTTCCAGGGGGCTGCTGGAATTTGGGAG	1542
Db	961	GGCCTGATACCAGAAACCTCGCATAGTCACTTCCAGGGGGCTGCTGGAATTTGGGAG	1020
Qy	1543	TCACACTGTGCGCAGGCAAGTGAAGAACCTTGGGGGAGAGGCGAGCTGTATGGGAGCCG	1602
Db	1021	TCACACTGTGCGCAGGCAAGTGAAGAACCTTGGGGGAGAGGCGAGCTGTATGGGAGCCG	1080
Qy	1603	TTTCAGAGCATGTCTTCTACTTTCAGCTGCGAGAGAGCTGGCCAGTCCAGAGTGTAGTCTC	1662
Db	1081	TTTCAGAGCATGTCTTCTACTTTCAGCTGCGAGAGAGCTGGCCAGTCCAGAGTGTAGTCTC	1140
Qy	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGGCTCCCATTTAGACAGATCCTGTCT	1722
Db	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGGCTCCCATTTAGACAGATCCTGTCT	1200
Qy	1723	AGGCCAGAGCGGCTCTTTCTATCTCGATGCTGTATGATGAGCCAGAGATGGGCTTTCAG	1782
Db	1201	AGGCCAGAGCGGCTCTTCTATCTCGATGCTGTATGATGAGCCAGAGATGGGCTTTCAG	1260

QY	1783	GAGCGAGATCTGAGCTCTGTCTGCACTGGAGCCAGCAACAGCGGGGATATGACCTGCTG	1842
Db	1261	GAGCCGAGTTCGTAGCTCTGTCTGCACTGGAGCCAGCAACAGCGGGGATATGACCTGCTG	1320
QY	1843	GGCAGTTTGTGCGGGGAAAATATATCTATCCGAGGCACTCTTCCGTATCAGCGCTCGAAC	1902
Db	1221	GGCAGTTTGTGCGGGGAAAATATATCTATCCGAGGCACTCTTCCGTATCAGCGCTCGAAC	1380
QY	1903	ACAGCTCTGCAAAACCTCATTTCTTCTTTGAGAGCGCACTTGGGTAGAGTCTGGGG	1962
Db	1381	ACAGCTCTGCAAAACCTCATTTCTTCTTTGAGAGCGCACTTGGGTAGAGTCTGGGG	1440
QY	1963	TTCTCTGAGTCCAGCAGAAAGSAAATATTTCTACAGATATTTCCACAGATGAAAGCAAGCA	2022
Db	1441	TTCTCTGAGTCCAGCAGAAAGSAAATATTTCTACAGATATTTCCACAGATGAAAGCAAGCA	1500
QY	2023	ATTAGAGCTTTAGATTGTTCATTAATCAACAAAGAGCTGTGGCCCTGTGTCTTTGTCCC	2082
Db	1501	ATTAGAGCTTTAGATTGTTCATTAATCAACAAAGAGCTGTGGCCCTGTGTCTTTGTCCC	1560
QY	2083	TGGGTGTCTTGCGCTGGCTGCACTTGGCTGATGAGAGATGAAAGCGGAAAGSAAAATC	2142
Db	1561	TGGGTGTCTTGCGCTGGCTGCACTTGGCTGATGAGAGATGAAAGCGGAAAGSAAAATC	1620
QY	2143	ACACTGACTTCCAGAGCAACCAACACCTCTGTATCATTAACCTTGGCCAGGCTCTCCAA	2202
Db	1621	ACACTGACTTCCAGAGCAACCAACACCTCTGTATCATTAACCTTGGCCAGGCTCTCCAA	1680
QY	2203	GCTCAGCCATTGGGAGCCCAAGCTCAGAGCCTGTCTCTGTGCGTGTAGAGGCATCTGG	2262
Db	1681	GCTCAGCCATTGGGAGCCCAAGCTCAGAGCCTGTCTCTGTGCGTGTAGAGGCATCTGG	1740
QY	2263	CAAAAAAAGACCCCTTTCAATCCAGCAGAGACCTCAGAGAAAGATGGGTATAGATGGGACCATC	2322
Db	1741	CAAAAAAAGACCCCTTTCAATCCAGCAGAGACCTCAGAGAAAGATGGGTATAGATGGGACCATC	1800
QY	2323	ATCTCCAACCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCCTGAGCTACAGC	2382
Db	1801	ATCTCCAACCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCCTGAGCTACAGC	1860
QY	2383	TTCAATTACCTCTGTCTTCCAAAGATTTCTTGGCAGCAATGTCTCTATGTCTTGGAGATAG	2442
Db	1861	TTCAATTACCTCTGTCTTCCAAAGATTTCTTGGCAGCAATGTCTCTATGTCTTGGAGATAG	1920
QY	2443	AAGGGGAGAGGTAAACATTTCTAATTGGCATATGATTTGGAAAAAGACGTAGAAAGCATAT	2502
Db	1921	AAGGGGAGAGGTAAACATTTCTAATTGGCATATGATTTGGAAAAAGACGTAGAAAGCATAT	1980
QY	2503	GGAATACATGGCCTGTTTGGGGCATCAACACAGTTTCTATTTGGCCTGTTTAATGTAT	2562
Db	1981	GGAATACATGGCCTGTTTGGGGCATCAACACAGTTTCTATTTGGCCTGTTTAATGTAT	2040
QY	2563	GAGGGGGAGAGAGATGAGAAACATTTTCACTGCGCGGTCTCAAGGGAGGAAACCTG	2622
Db	2041	GAGGGGGAGAGAGATGAGAAACATTTTCACTGCGCGGTCTCAAGGGAGGAAACCTG	2100
QY	2623	ATGCAATGGGTCCCGTCCCTGCAAGCTGCTGTCAGCACAATCTCTGAGTCCCTCAC	2682
Db	2101	ATGCAATGGGTCCCGTCCCTGCAAGCTGCTGTCAGCACAATCTCTGAGTCCCTCAC	2160
QY	2683	TGCTTTGATCCAGACTTCGGAAACAAACGTTTCTCTGACACAAGTGAATGGCCATTTTGAAGAA	2742
Db	2161	TGCTTTGATCCAGACTTCGGAAACAAACGTTTCTCTGACACAAGTGAATGGCCATTTTGAAGAA	2220
QY	2743	ATGGGCAATGTGTGTAAGAAACAGACATGAACTCTTAATGTGCACTTCTCTCATTAATTC	2802
Db	2221	ATGGGCAATGTGTGTAAGAAACAGACATGAACTCTTAATGTGCACTTCTCTCATTAATTC	2280
QY	2803	AGCCGCCACGTGAAGAAAGCTTCAAGCTGATTTGAGGGCAGGCAACAGATCAACATGAGAC	2862
Db	2281	AGCCGCCACGTGAAGAAAGCTTCAAGCTGATTTGAGGGCAGGCAACAGATCAACATGAGAC	2340

QY 2863 CCCACATGAGTCTGTTCAAGTGGATCCAGTCAAGATGCTATTGGAGATTCTC 2922
 DB 2341 CCCACATGAGTCTGTTCAAGTGGATCCAGTCAAGATGCTATTGGAGATTCTC 2400
 QY 2923 TTCTCCGCTCTCAAGGTCAACCAAACTTGAAGAGCTGGACTTAAGTGAATCTGCTG 2982
 DB 2401 TTCTCCGCTCTCAAGGTCAACCAAACTTGAAGAGCTGGACTTAAGTGAATCTGCTG 2460
 QY 2983 AGCACACTGACAGTGAAGAGCTTTTGAAGCCTTGAAGCGCTCCGCTGCTCTGAG 3042
 DB 2461 AGCACACTGACAGTGAAGAGCTTTTGAAGCCTTGAAGCGCTCCGCTGCTCTGAG 2520
 QY 3043 ACCCTGGGTTGGCTGCTGCTGCTGCTCAAGCTGAGACTGCAAGACTTGGCTTGGG 3102
 DB 2521 ACCCTGGGTTGGCTGCTGCTGCTGCTCAAGCTGAGACTGCAAGACTTGGCTTGGG 2580
 QY 3103 CTGAGAGCCAAACAGACCTTGAACCTGAGCTGGAAGCTTAAGTGTCTCAAGATGCT 3162
 DB 2581 CTGAGAGCCAAACAGACCTTGAACCTGAGCTGGAAGCTTAAGTGTCTCAAGATGCT 2640
 QY 3163 GGAAGCCAAACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTCAAGCACTGAG 3222
 DB 2641 GGAAGCCAAACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTCAAGCACTGAG 2700
 QY 3223 CTGGTCAAGCTGAGCTCAAGCTGAGACTGAGACTGAGACTGAGACTGAGACTGAG 3282
 DB 2701 CTGGTCAAGCTGAGCTCAAGCTGAGACTGAGACTGAGACTGAGACTGAGACTGAG 2760
 QY 3283 AGCCCAAGCTGAGAGAGTGAACCTGAGAGCAAACTTAAGTGAAGCTTGGCTGAG 3342
 DB 2761 AGCCCAAGCTGAGAGAGTGAACCTGAGAGCAAACTTAAGTGAAGCTTGGCTGAG 2820
 QY 3343 CTGCTCTGAGAGGGGCTCAAGGACTCTGCTGCAAACTCAAGCTGAGAGGGCTGAGAG 3402
 DB 2821 CTGCTCTGAGAGGGGCTCAAGGACTCTGCTGCAAACTCAAGCTGAGAGGGCTGAGAG 2880
 QY 3403 ACAACTCTGAGAGTGAAGTGAAGGAGAGAGCTGAGAGGGGCTGAGAGAGAACTCAG 3462
 DB 2881 ACAACTCTGAGAGTGAAGTGAAGGAGAGAGCTGAGAGGGGCTGAGAGAGAACTCAG 2940
 QY 3463 CTGCTCACTTCAAGCAGAGAGAACTCAAGTGAAGAGCTTGAAGAGGGCTGAGATAG 3522
 DB 2941 CTGCTCACTTCAAGCAGAGAGAACTCAAGTGAAGAGCTTGAAGAGGGCTGAGATAG 3000
 QY 3523 GGAAGATGAGATATAGACATCTCACTCAAGGGGAGAGACTGAGATCAAGAGGGGCTG 3582
 DB 3001 GGAAGATGAGATATAGACATCTCACTCAAGGGGAGAGACTGAGATCAAGAGGGGCTG 3060
 QY 3583 GCTTCCCATGTTGCTCAAGGCTAATCTCAAACTCCCTGAGAGAGAGATCTTCCCAAT 3642
 DB 3061 GCTTCCCATGTTGCTCAAGGCTAATCTCAAACTCCCTGAGAGAGAGATCTTCCCAAT 3120
 QY 3643 GCTGAGATGAGAGAGAAAGCTCCCAAGAGTGAAGTCCGATGGAATCTTGTGCTGCT 3702
 DB 3121 GCTGAGATGAGAGAGAAAGCTCCCAAGAGTGAAGTCCGATGGAATCTTGTGCTGCT 3180
 QY 3703 TCTCTGCTCTCAAGGGGAGCTTGAAGAGCTTTGGGAGCTGAGAGATCTTCTGG 3762
 DB 3181 TCTCTGCTCTCAAGGGGAGCTTGAAGAGCTTTGGGAGCTGAGAGATCTTCTGG 3240
 QY 3763 GGGCCCAAGGGGGCTGAGGCTAAGAGTGAAGTGAAGAAAGAACTTGAAGAGTT 3822
 DB 3241 GGGCCCAAGGGGGCTGAGGCTAAGAGTGAAGTGAAGAAAGAACTTGAAGAGTT 3300
 QY 3823 CACTTCCCTGAGTGGCTCTCAAGCTGAGGCTCAAGAGGGTCTCTGCTTGTGATGAGA 3882
 DB 3301 CACTTCCCTGAGTGGCTCTCAAGCTGAGGCTCAAGAGGGTCTCTGCTTGTGATGAGA 3360
 QY 3883 GAAAGGCTGAGAGCTTGAAGATTTCTGTGTGTGAGACAGCTTCTGGGTGAGATCAAC 3942
 DB 3361 GAAAGGCTGAGAGCTTGAAGATTTCTGTGTGTGAGACAGCTTCTGGGTGAGATCAAC 3420
 QY 3943 CCAAGCAGAGCTGAGTGGAGAGGGGCTCTGCTGAGATCAAGAGCTGAGAGCTGAGACT 4002

DB 3421 CCACAGACAGCTGAGATGAGTGGAGAGGGCTCTGCTGAGATCAAGGCTGAGAGCT 3480
 QY 4003 GTGAGAGCTGAGACTCTCCCTCACTTTGGCTCTCCAGAGGGGCTCAATGAGACATCC 4062
 DB 3481 GTGAGAGCTGAGACTCTCCCTCACTTTGGCTCTCCAGAGGGGCTCAATGAGACATCC 3540
 QY 4063 CTGTTCCAAATGGCCCACTTAAAGAGAGAGAGTCTCTGAGAGAGCCAGAGGAGT 4122
 DB 3541 CTGTTCCAAATGGCCCACTTAAAGAGAGAGAGTCTCTGAGAGAGCCAGAGGAGT 3600
 QY 4123 GAGCTGATCAATAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4182
 DB 3601 GAGCTGATCAATAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
 QY 4183 ATGATCCATATAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4242
 DB 3661 ATGATCCATATAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
 QY 4243 GTCCATCTGAG 4302
 DB 3721 GTCCATCTGAG 3780
 QY 4303 ----- 4302
 DB 3781 AAGGCCATAGATGATCTAGAAATGAATTCAGATTGTGAGAGATCCAAAGCACC 3840
 QY 4303 ----- 4302
 DB 3841 CTGACCCCACTTATATAGGCTGCTGTTACATGCTGTCTGAGTCTGAGAGAGATCTG 3900
 QY 4303 -----AAGAACTGAGAGCTGCTGTTATCGAAGAGAGAGAGAGAGAGAGAGAG 4350
 DB 3901 GAAATACTCCCAAG 3960
 QY 4351 TCGAGATTCTAGCTGAG 4410
 DB 3961 TCGAGATTCTAGCTGAG 4020
 QY 4411 GATGAGACTGCTGAG 4470
 DB 4021 GATGAGACTGCTGAG 4080
 QY 4471 CTGATCCCTGAG 4530
 DB 4081 CTGATCCCTGAG 4140
 QY 4531 TTGTGAGACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4590
 DB 4141 TTGTGAGACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4200
 QY 4591 GACAACTGATGAG 4650
 DB 4201 GACAACTGATGAG 4260
 QY 4651 ACGAGGCTCAG 4710
 DB 4261 ACGAGGCTCAG 4320
 QY 4711 AAAGATGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4770
 DB 4321 AAAGATGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380
 QY 4771 GAG 4812
 DB 4381 GAG 4422

RESULT 10
 AF229061 AF229061 4200 bp mRNA linear PRI 02-FEB-2001
 LOCUS Homo sapiens NAC-gamma splice variant (NAC) mRNA, complete cds,
 DEFINITION alternatively spliced.

ACCESSION AF229061.1 GI:12656108
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE AUTHORS Chu, Z.-L., Pio, F., Xie, Z., Godzik, A. and Reed, J.C.
 TITLE NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c pathway for apoptosis
 JOURNAL 2 (bases 1 to 4200)
 REFERENCE AUTHORS Chu, Z.-L. and Reed, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
 FEATURES
 source
 1..4200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /cell_type="activated Jurkat T cells"
 1..4200
 /gene="NAC"
 1..4200
 /gene="NAC"
 /note="Apaf-1/Ced-4 family apoptosis regulator; contains NB-ARC domain, CARD domain, and leucine-rich repeats (LRR)"
 /codon_start=1
 /product="NAC-gamma splice variant"
 /protein_id="AAK00750.1"
 /db_xref="GI:12656109"
 /translation="MAGGARGRLACYLEFKKEELKEFOLLANKASHRSSEGETPAQ
 PERTSGNEVASLYVAQYGEORAWDLAHTHEWGLKRSICAOAGASHSPFPSPES
 PHLGSPQPTSTAVLMPWIHELPAQCTGSSRRVLRQLPTSGRMERISASLILYOL
 PSSEHDHSPSOESPNAPTSTAVLGSWSPPPSLAPBOEAPGTOMPDENSGIYTER
 IRRERREKSEKGRPMMAVVGTPPOAHTSLQPHHPMPVSRELSCTWPKNDPNQ
 KPTQLDLPHPRSODPLVRSMDPYEENRGILRLFGGLQLOBERIYLIGLO
 AAGIGKSTLAROYEMAGRGQLYEDRPOHVSREGLRELSAQSKYVSLAEILCKDTAP
 APIHQILSRPERLLFILDGVDPEGVWLPSESELCLMSOPQADALLGLNLTILP
 EASLILARTTALQNLPSLEQAWRVYLGESESREKEYFRYPTDRQALRAVLK
 SNEKLMAELCVPMWSWLACTLMQOMRKKEKLTSTKTTTLCLHYLAQAPLIGS
 QLRDLGLAAGIGWCKTLFSPDDLRKGLDGLIISFLKGILOEHPILASYSFIH
 CPOFFPAMASVYLEDEKRGKSHNCITIDKTLTAAGIGHGFGASTTRFLIGLSDSG
 EREHENTFHCHLSQGRINLMQVPSLQILLOPHSLESILHCLYETRNKPTFLTOYMAHFE
 MWQVEITDMELLVCTFCIKESRHVKQLLEGHRHSITWSCTMYVLRFWCVTDAVQ
 ILFEVLVLRNRLKELDLSGNSLSHSAVSLKXTLRPRCLLETLRAGLCCATDCDK
 LAFILRANQUTLELIDLFNNVLTDAAGKLCQLRLQSPCKQLRLQVSGCCDCCD
 ASVLSASPSKLELDLQONLDDVGRLLCELRHPACKLILRGKSWTPTFGDGTGS
 MNSNTSLIKRORLGSERASVHAQANIKLDVSKTIPRLAIESSSPVVVELLCYB
 SPASQGLHTPEPLGTDDEFWGPTPEPVATVEYVDKKNLYRHPFAGSTRPENTLCYV
 MREAVTVEIEFCVWDQFLGELINPQHSMMVAPSLDLIDKAPGAVVAHLPHVLAQGG
 VDSILFQWAKHEKGMLEKPARVELHITVLENSFSPLGLIMHINALFIVTSG
 VILVHRYAPEVTEFHYLIIPDSGISIRKELICTYSGPGDGLSFYVGLSGSGRLQV
 KDKDELTMVAIKYLPGLMIPATTLIPARLIAPSPDAPQLFLHVOYRQLTARVLT
 SVEVLDLKLHGVTSQSDYREVLADNTRPSQMRKLFSLSGSDMKCKDGLYLAKETHH
 PHLMELMEKSKKGLPLPS"

Db	61	CTGAAGAGTTCCAGCTTCTGCTCGCCAAATTAAGCGCACTCCAGAGACTCTTCGGGTGAG	120
Qy	643	ACAACCGCTCAGCCAAGAAAGACGATGCGATGAGTGGCTTCGTACTGTGTGGCTCAG	702
Db	121	ACACCGCTCAGCCAGAGAAAGACGATGCGATGAGTGGCTTCGTACTGTGTGGCTCAG	180
Qy	703	TATGGGGAGCAGGGGCTGTGGGACTAGAGCTCCTCATCTGTGGGAGCAGATGGGGCTGAG	762
Db	181	TATGGGGAGCAGGGGCTGTGGGACTAGAGCTCCTCATCTGTGGGAGCAGATGGGGCTGAG	240
Qy	763	TCACGTGCGCCCAAGGCCAGAGGGGACAGGCCACTCTCCCTCATCTCCCTCAGGCCA	822
Db	241	TCACGTGCGCCCAAGGCCAGAGGGGACAGGCCACTCTCCCTCATCTCCCTCAGGCCA	300
Qy	823	AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTTCAACCGCATGTCTAATGCCCTGG	882
Db	301	AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTTCAACCGCATGTCTAATGCCCTGG	360
Qy	883	ATCCATGAATTTGGCGGCGGGGTGCACTCCAGGCTCAGAGAGAAAGGTTTGAACAACCTG	942
Db	361	ATCCATGAATTTGGCGGCGGGGTGCACTCCAGGCTCAGAGAGAAAGGTTTGAACAACCTG	420
Qy	943	CTGACACATCTGAGACGCGCTGAGAGAAATCTGTGCTCATCTCCCTCFACCAAGCTT	1002
Db	421	CTGACACATCTGAGACGCGCTGAGAGAAATCTGTGCTCATCTCCCTCFACCAAGCTT	480
Qy	1003	CCAAGCTCCCAAGCCATGAGTCTCCAGCCAGAGTCAACCAACGCCCCCATCTCACA	1062
Db	481	CCAAGCTCCCAAGCCATGAGTCTCCAGCCAGAGTCAACCAACGCCCCCATCTCACA	540
Qy	1063	GCAGTGTGTGGGAGCTGTGGGATCCCCACTCAGCCCAAGCTTAGCAACCCAGAGCAAGAG	1122
Db	541	GCAGTGTGTGGGAGCTGTGGGATCCCCACTCAGCCCAAGCTTAGCAACCCAGAGCAAGAG	600
Qy	1123	GCCTCTGGGAGCCCAATGGCTCTGTGATGAACGTCAGAAATTTACTACAGAAATCAGA	1182
Db	601	GCCTCTGGGAGCCCAATGGCTCTGTGATGAACGTCAGAAATTTACTACAGAAATCAGA	660
Qy	1183	GAAAGAGAGAGAGAAATCAGAGAAAGGACGAGCCCCATGTGGGACGGGTGTGAGAACG	1242
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGGACGAGCCCCATGTGGGACGGGTGTGAGAACG	720
Qy	1243	CCCCCACAAGCGCACACCAAGCTTACAGCCCCACACACCCATGTGGAGCTTCTGTGAGA	1302
Db	721	CCCCCACAAGCGCACACCAAGCTTACAGCCCCACACACCCATGTGGAGCTTCTGTGAGA	780
Qy	1303	GAGAGCTCTGTCCATGAGGCTGTGGAAATGAGGATTTTAAACAAAATTCACACAG	1362
Db	781	GAGAGCTCTGTCCATGAGGCTGTGGAAATGAGGATTTTAAACAAAATTCACACAG	840
Qy	1363	CTGCTACTTTACAAAGACTCACCAGAGGCAAGATCCCTGTCAAGAGAGCTGG	1422
Db	841	CTGCTACTTTACAAAGACTCACCAGAGGCAAGATCCCTGTCAAGAGAGCTGG	900
Qy	1423	CTGATTAATGTGAGAGGAATCAGAGCATTTAATGAGATCAGAGACTTAATTTGGCCCA	1482
Db	901	CTGATTAATGTGAGAGGAATCAGAGCATTTAATGAGATCAGAGACTTAATTTGGCCCA	960
Qy	1483	GGCTGTGATTCACCAAGAACTCCCATGTGTACTGTGAGGGGGGTGTGGAAATTTGGGAA	1542
Db	961	GGCTGTGATTCACCAAGAACTCCCATGTGTACTGTGAGGGGGGTGTGGAAATTTGGGAA	1020
Qy	1543	TCAACACTGGCCAGGCGAGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTATGGGAAACGC	1602
Db	1021	TCAACACTGGCCAGGCGAGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTATGGGAAACGC	1080
Qy	1603	TTCCAGCATGTCTTTACTTACAGTGTGACAGAGCTGGCCAGCTCAGACAGGTGTGATGTCT	1662
Db	1081	TTCCAGCATGTCTTTACTTACAGTGTGACAGAGCTGGCCAGCTCAGACAGGTGTGATGTCT	1140
Qy	1663	GTGAGCTCATTCGGAAGATGGGACAGGCACTCCGGCTCCCATTAACAGATCTGTCT	1722

Db 1141 GCTGAGCTCATCGGAAAAGATGGGACAGCACTCCGGCTCCATTAGACAGATCTGTCT 1200
Qy 1723 AGGCGAAGAGGGCTGCTCTTCAATCTCGATGATGATAGGAGGAGGATGGGTCTTTCAG 1782
Db 1201 AGGCGAAGAGGGCTGCTCTTCAATCTCGATGATGATAGGAGGAGGATGGGTCTTTCAG 1260
Qy 1783 GAGCCGAGTCTGAGCTCTGTCTGACAGTGGAGCCAGCCAGCCGGGAGATGACCTGCTG 1842
Db 1261 GAGCCGAGTCTGAGCTCTGTCTGACAGTGGAGCCAGCCAGCCGGGAGATGACCTGCTG 1320
Qy 1843 GGCAGTTTGTGGGAAAACATACTTCCGAGGCACTCTTCTGATCAGCGCTCGAGCC 1902
Db 1321 GGCAGTTTGTGGGAAAACATACTTCCGAGGCACTCTTCTGATCAGCGCTCGAGCC 1380
Qy 1903 ACAGCTCTGAGAACTCTCATCTCTTCTTGGAGAGGAGCGTTGGGTAGAGTCTTGGG 1962
Db 1381 ACAGCTCTGAGAACTCTCATCTCTTCTTGGAGAGGAGCGTTGGGTAGAGTCTTGGG 1440
Qy 1963 TTCTCTGAGTCCAGAGGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGAGGAAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 1500
Qy 2023 ATTAGAGCTTTAGGTTGGTCAAACTAAACAAAGAGCTTGGGCTGTGTCTTGTGCC 2082
Db 1501 ATTAGAGCTTTAGGTTGGTCAAACTAAACAAAGAGCTTGGGCTGTGTCTTGTGCC 1560
Qy 2083 TGGGTGTCTGTGGCTGCTGCACTTGTCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGG 2142
Db 1561 TGGGTGTCTGTGGCTGCTGCACTTGTCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Qy 2143 ACACTGACTTCAAGAGCAACCAACCACTGTCTGATTAATCTTGGCCAGGCTCTGCCA 2202
Db 1621 ACACTGACTTCAAGAGCAACCAACCACTGTCTGATTAATCTTGGCCAGGCTCTGCCA 1680
Qy 2203 GCTCAGCATTGGGAGCCCACTGAGAGAGCTGTCTCTGAGCTGAGAGGAGCTGAG 2262
Db 1681 GCTCAGCATTGGGAGCCCACTGAGAGAGCTGTCTCTGAGCTGAGAGGAGCTGAG 1740
Qy 2263 CAAAAAAGAGCCCTTTCAGTCCAGATGAGCTCAGAGAGGAGTGGGTAGATGGGAGGAG 2322
Db 1741 CAAAAAAGAGCCCTTTCAGTCCAGATGAGCTCAGAGAGGAGTGGGTAGATGGGAGGAG 1800
Qy 2323 ATCTCCACTTCTTGAAGATGGGTATCTTCAAGAGAGCCCACTGAGCTGAGCTAGC 2382
Db 1801 ATCTCCACTTCTTGAAGATGGGTATCTTCAAGAGAGCCCACTGAGCTGAGCTAGC 1860
Qy 2383 TTCAATTCACCTGTCTTCAAGAGTCTTGGAGCAATGTCTTAATGTCTTGGAGATGAG 2442
Db 1861 TTCAATTCACCTGTCTTCAAGAGTCTTGGAGCAATGTCTTAATGTCTTGGAGATGAG 1920
Qy 2443 AAGGAGAGAGTAAACATTTCTAATTCATCATAGATTTGAGAAAGAGCTAGAGCATAT 2502
Db 1921 AAGGAGAGAGTAAACATTTCTAATTCATCATAGATTTGAGAAAGAGCTAGAGCATAT 1980
Qy 2503 GGAATATCATGGCCGTGTGGGAGCATCAACCAAGTTCCTAATTTGGGCTGTTAAGGAT 2562
Db 1981 GGAATATCATGGCCGTGTGGGAGCATCAACCAAGTTCCTAATTTGGGCTGTTAAGGAT 2040
Qy 2563 GAGGAGAGAGAGATGAGAGACATCTTTCAGTCCGAGCTGTCTCAGGAGAGGAGACCTG 2622
Db 2041 GAGGAGAGAGAGATGAGAGACATCTTTCAGTCCGAGCTGTCTCAGGAGAGGAGACCTG 2100
Qy 2623 ATGCAATGGGTCTCCCTGCACTGCTGTGTGAGAGCACTCTCTGAGAGTCCCTTCAC 2682
Db 2101 ATGCAATGGGTCTCCCTGCACTGCTGTGTGAGAGCACTCTCTGAGAGTCCCTTCAC 2160
Qy 2683 TGCTTGAACAGACTCGGAAACAAACGTTCTGACACAGAGTATGGCCATTTTCAGAGAA 2742
Db 2161 TGCTTGAACAGACTCGGAAACAAACGTTCTGACACAGAGTATGGCCATTTTCAGAGAA 2220
Qy 2743 ATGGGCAATGTGTAGAAACAGACATGAGCTTTAGTGTGCACTTTCTGCAATTAATTC 2802
Db 2221 ATGGGCAATGTGTAGAAACAGACATGAGCTTTAGTGTGCACTTTCTGCAATTAATTC 2280

Qy 2803 AGCCGCAAGTGAAGAGCTTCAAGTATGAGGAGGAGGAGAGACAGATGCAATGAGGAGC 2862
Db 2281 AGCCGCAAGTGAAGAGCTTCAAGTATGAGGAGGAGGAGAGACAGATGCAATGAGGAGC 2340
Qy 2863 CCACCAATGATCTCTGTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGAGATTTCTC 2922
Db 2341 CCACCAATGATCTCTGTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGAGATTTCTC 2400
Qy 2923 TTCTCCGTCTCAAGGTCAACAGAAACCTGAAGAGCTGAGCCTTAAGTGAACCTGCTG 2982
Db 2401 TTCTCCGTCTCAAGGTCAACAGAAACCTGAAGAGCTGAGCCTTAAGTGAACCTGCTG 2460
Qy 2983 AGCACTCTGCAATGAAGATCTTTGTAAGACCTTGAAGACCCCTGCTGCTGCTGAG 3042
Db 2461 AGCACTCTGCAATGAAGATCTTTGTAAGACCTTGAAGACCCCTGCTGCTGCTGAG 2520
Qy 3043 ACCCTGGGTTGGCTGAGCTGAGCTCAAGCTGAGGAGCTGCAAGGAGCTTGTGCTTGGG 3102
Db 2521 ACCCTGGGTTGGCTGAGCTGAGCTGAGCTGAGGAGCTGCAAGGAGCTTGTGCTTGGG 2580
Qy 3103 CTGAGAGCCAAACCAACCTTGAACGAGCTGAGCCTGAGCTTCAATGTCTCACGAGTGT 3162
Db 2581 CTGAGAGCCAAACCAACCTTGAACGAGCTGAGCCTGAGCTTCAATGTCTCACGAGTGT 2640
Qy 3163 GAGCCAAACCACTTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTGCAAGCACTGAG 3222
Db 2641 GAGCCAAACCACTTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTGCAAGCACTGAG 2700
Qy 3223 CTGGTCAAGTGTGGCTCAAGCTGAGCTGAGCCTGAGCCTGCTGCTTATGAGGCT 3282
Db 2701 CTGGTCAAGTGTGGCTCAAGCTGAGCTGAGCCTGAGCCTGCTGCTTATGAGGCT 2760
Qy 3283 AGCCCAAGCTGAGAGAGCTGAGCCTGAGCAGAGCAACCTGAGTGAAGTGGCTGAGCA 3342
Db 2761 AGCCCAAGCTGAGAGAGCTGAGCCTGAGCAGAGCAACCTGAGTGAAGTGGCTGAGCA 2820
Qy 3343 CTGGCTGAGAGGAGCTGAGAGCTGCTGCTGAGCACTGAGCCTGAGTGAAGTGGCTGAGCA 3402
Db 2821 CTGGCTGAGAGGAGCTGAGAGCTGCTGCTGAGCACTGAGCCTGAGTGAAGTGGCTGAG 2860
Qy 3403 ACAACTGAGTATGATGAGTGAAGGAGCAAGCTGAGGAGCCTGAGAGAGGAAACCTGAG 3462
Db 2870 ----- 2860
Qy 3463 CTGCTCATCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3522
Db 2870 ----- GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2910
Qy 3523 GAGGAGTGAATTAAGACATCTCTCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3582
Db 2911 GAGGAGTGAATTAAGACATCTCTCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2970
Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 3642
Db 2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 3030
Qy 3643 GCTGAGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3702
Db 3031 GCTGAGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3090
Qy 3703 TCTTCCGTCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3762
Db 3091 TCTTCCGTCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3150
Qy 3763 GAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3822
Db 3151 GAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3210
Qy 3823 CACTTCCGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3882
Db 3211 CACTTCCGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3270

QY	3883	GAACCGGTGAACCGTTGAGATTGAATTCGTGTGTGGGACACAGTTCTCTGGAGATCAAC	3942
Dp	3271	GAACCGGTGAACCGTTGAGATTGAATTCGTGTGTGGGACACAGTTCTCTGGAGATCAAC	3330
QY	3943	CCACAGCAACAGCTGAGTGTGGCAGGGGCTCTGTGGAATCAAGAGCTGAGCTGGAAGCT	4002
Dp	3331	CCACAGCAACAGCTGAGTGTGGCAGGGGCTCTGTGGAATCAAGAGCTGAGCTGGAAGCT	3390
QY	4003	GTGGAAGCTGTGACATCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGACATATCC	4062
Dp	3391	GTGGAAGCTGTGACATCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGACATATCC	3450
QY	4063	CTGTTCCTCAAAATGGGCCACTTTTAAAGAGAGGGGAGATGCTCTGAGAAAGCGACCGAGGTTG	4122
Dp	3451	CTGTTCCTCAAAATGGGCCACTTTTAAAGAGAGGGGAGATGCTCTGAGAAAGCGACCGAGGTTG	3510
QY	4123	GAGCTGCATCAATAGTTCTGTGAAAAACCCACAGTTCTCCCTCTGGAGAGTCCCTCTGAAA	4182
Dp	3511	GAGCTGCATCAATAGTTCTGTGAAAAACCCACAGTTCTCCCTCTGGAGAGTCCCTCTGAAA	3570
QY	4183	ATGATTCATTAATGCCCTGCGCTTCATTTCCTCGTCACTCTGTGTGTGTTGCTTTACACCGC	4242
Dp	3571	ATGATTCATTAATGCCCTGCGCTTCATTTCCTCGTCACTCTGTGTGTGTTGCTTTACACCGC	3630
QY	4243	GTTCATCCTGTGAGAAAGTCACTTCCACTCTACCTGATCCCAAGGATCGATGCTCCATTCGG	4302
Dp	3631	GTTCATCCTGTGAGAAAGTCACTTCCACTCTACCTGATCCCAAGGATGCTGCTCCATTCGG	3690
QY	4303	AAGGAATGAGAGCTCTGCTATTCGAAAGCCCTGTGAGAAAGACAGAGTTCTCCGAGTTCTAC	4362
Dp	3691	AAGGAATGAGAGCTCTGCTATTCGAAAGCCCTGTGAGAAAGACAGAGTTCTCCGAGTTCTAC	3750
QY	4363	GTTGGCCACTTTGGGATCAAGGATCAAGCTGCAAGTGAAGACAGAAAGATGAGACTCTTG	4422
Dp	3751	GTTGGCCACTTTGGGATCAAGGATCAAGCTGCAAGTGAAGACAGAAAGATGAGACTCTTG	3810
QY	4423	GTGTGGGAGGCGTTTGGTGAACCAAGAGATCTCATGCGCTGGAATCACTCTGATCCCTCA	4482
Dp	3811	GTGTGGGAGGCGTTTGGTGAACCAAGAGATCTCATGCGCTGGAATCACTCTGATCCCTCA	3870
QY	4483	GCCCGCATATGCCGTAACCTTCACTCTGTGATGCCCGCAGTTGCTGCACTTTGTGGACAG	4542
Dp	3871	GCCCGCATATGCCGTAACCTTCACTCTGTGATGCCCGCAGTTGCTGCACTTTGTGGACAG	3930
QY	4543	TATCGAAGACAGCTGATATGCCGAGTGCATTCGGTGGAGGTTGTCTTGGACAAATCTGCAT	4602
Dp	3931	TATCGAAGACAGCTGATATGCCGAGTGCATTCGGTGGAGGTTGTCTTGGACAAATCTGCAT	3990
QY	4603	GGAAGAGGTCGTGAACCAAGGACAGATGACAGAGGGGTGCTGTGAGAAACAAGAGCCCAAGC	4662
Dp	3991	GGAAGAGGTCGTGAACCAAGGACAGATGACAGAGGGGTGCTGTGAGAAACAAGAGCCCAAGC	4050
QY	4663	CAGATGGGAAAGCTGTTGAGCTTGGAGCAGATCTTGGAGCCGGAAGTGCAGAAAGATGGAATC	4722
Dp	4051	CAGATGGGAAAGCTGTTGAGCTTGGAGCAGATCTTGGAGCCGGAAGTGCAGAAAGATGGAATC	4110
QY	4723	TACCAAGCTCTGAAGAGAACCCATCTCACTTATGAAACTCTGAGAAAGGGCAGC	4782
Dp	4111	TACCAAGCTCTGAAGAGAACCCATCTCACTTATGAAACTCTGAGAAAGGGCAGC	4170
QY	4783	AAAAAGGACCTCTGCACTCAAGCAGCTGA	4812
Dp	4171	AAAAAGGACCTCTGCACTCAAGCAGCTGA	4200

[illegible]

QY 883 ATCCATGATTTGCGCGGGGTCACCCAGGGCTCAGAGAGAAAGGTTTGAACAGCTG 942
| | | | |
Db 361 ATCCATGATTTGCGCGGGGTCACCCAGGGCTCAGAGAGAAAGGTTTGAACAGCTG 420
| | | | |
QY 943 CTTGACACATCTGAGCGCCGCTGAGAGAAATCTCTGCTCTACCTCTCTACCAAGCTCTT 1002
| | | | |
Db 421 CTTGACACATCTGAGCGCCGCTGAGAGAAATCTCTGCTCTACCTCTCTACCAAGCTCTT 480
| | | | |
QY 1003 CCAAGCTCCCCAGACCTAGAGTCTCAAGGCAAGAGTCACCCAAGCCCCCATCTCA 1062
| | | | |
Db 481 CCAAGCTCCCCAGACCTAGAGTCTCAAGGCAAGAGTCACCCAAGCCCCCATCTCA 540
| | | | |
QY 1063 GCAGTGCTGGGAGCTGGGAGTCCCACTCAGGCCAGCCCTAGGACCCAGAGCAGAG 1122
| | | | |
Db 541 GCAGTGCTGGGAGCTGGGAGTCCCACTCAGGCCAGCCCTAGGACCCAGAGCAGAG 600
| | | | |
QY 1123 GCTCTTGGGACCCAAATGCTCTGATGAAACGTCAAGAAATTTACTACAGAAATCAGA 1182
| | | | |
Db 601 GCTCTTGGGACCCAAATGCTCTGATGAAACGTCAAGAAATTTACTACAGAAATCAGA 660
| | | | |
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGAGGCGGCTGTGAGAACG 1242
| | | | |
Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGAGGCGGCTGTGAGAACG 720
| | | | |
QY 1243 CCCCCAGAGGCGACACAGGCTTACAGCCCAACCAACCATGAGGAGCCCTGTGAGA 1302
| | | | |
Db 721 CCCCCAGAGGCGACACAGGCTTACAGCCCAACCAACCATGAGGAGCCCTGTGAGA 780
| | | | |
QY 1303 GAGAGCTCTGTTTCCATGAGGCTCTGAGAAATAGAGATTTTAAACAAAATTCACAG 1362
| | | | |
Db 781 GAGAGCTCTGTTTCCATGAGGCTCTGAGAAATAGAGATTTTAAACAAAATTCACAG 840
| | | | |
QY 1363 CTGCTACTTTTCAAAAGACTCACCACCAAGAACCAAGATCCCTGCTGTCAGAGAGCTG 1422
| | | | |
Db 841 CTGCTACTTTTCAAAAGACTCACCACCAAGAACCAAGATCCCTGCTGTCAGAGAGCTG 900
| | | | |
QY 1423 CTTGATTAATGAGAGAGAAATCAGAGCACTTTAATTGATCAGAGACTTAATTGGGCCA 1482
| | | | |
Db 901 CTTGATTAATGAGAGAGAAATCAGAGCACTTTAATTGATCAGAGACTTAATTGGGCCA 960
| | | | |
QY 1483 GGGCTGAGTACCCAGAACTCTGATAGTCACTGAGGAGGCTGCTGAGAAATTTGAGAG 1542
| | | | |
Db 961 GGGCTGAGTACCCAGAACTCTGATAGTCACTGAGGAGGCTGCTGAGAAATTTGAGAG 1020
| | | | |
QY 1543 TCAACACTGCGCAGGCGAGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTATGAGGAGCCG 1602
| | | | |
Db 1021 TCAACACTGCGCAGGCGAGTGAAGAAAGCTGGGGGAGAGGCCAGCTGTATGAGGAGCCG 1080
| | | | |
QY 1603 TTCCAGCATGTCTTCTACTTCAAGCTGAGAGAGTGGCCAGTCCAGAGTGTAGAGCTC 1662
| | | | |
Db 1081 TTCCAGCATGTCTTCTACTTCAAGCTGAGAGAGTGGCCAGTCCAGAGTGTAGAGCTC 1140
| | | | |
QY 1663 GCTGAGCTCATCGGAAAGATGGAGACAGCACTCCGCTCCCATTTAGACAGATCTGCT 1722
| | | | |
Db 1141 GCTGAGCTCATCGGAAAGATGGAGACAGCACTCCGCTCCCATTTAGACAGATCTGCT 1200
| | | | |
QY 1723 AGGCGAGAGCGGCTGCTCTTCAATCTCTGATGTGTATGATGAGCCAGATGGGCTTTCAG 1782
| | | | |
Db 1201 AGGCGAGAGCGGCTGCTCTTCAATCTCTGATGTGTATGATGAGCCAGATGGGCTTTCAG 1260
| | | | |
QY 1783 GAGCGAGTCTGAGCTGTGCTGACCTGAGCCAGCAAGCCGAGGAGTGTGACCTG 1842
| | | | |
Db 1261 GAGCGAGTCTGAGCTGTGCTGACCTGAGCCAGCAAGCCGAGGAGTGTGACCTG 1320
| | | | |
QY 1843 GGCAGTTTGTGGGAAACCTACTTCCGAGGACATCTTCTGATCAGCGCTCGAGCC 1902
| | | | |
Db 1321 GGCAGTTTGTGGGAAACCTACTTCCGAGGACATCTTCTGATCAGCGCTCGAGCC 1380
| | | | |
QY 1903 ACAGCTCTGAGAAACCTCATTTCTTTTGAAGAGGACGTTGGGTAGAGGCTCTGGG 1962
| | | | |
Db 1381 ACAGCTCTGAGAAACCTCATTTCTTTTGAAGAGGACGTTGGGTAGAGGCTCTGGG 1440
| | | | |

QY 1963 TTCTGAGTCCAGCAGAGAAAGATATTTCTACAGATATTTCAACAGATGAAGGCAAGCA 2022
| | | | |
Db 1441 TTCTGAGTCCAGCAGAGAAAGATATTTCTACAGATATTTCAACAGATGAAGGCAAGCA 1500
| | | | |
QY 2023 ATTAGAGCTTTAGGTTGCTCAATCAAAAGAGCTGAGGCTGCTGCTTTGCTCC 2082
| | | | |
Db 1501 ATTAGAGCTTTAGGTTGCTCAATCAAAAGAGCTGAGGCTGCTGCTTTGCTCC 1560
| | | | |
QY 2083 TGGGTGCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2142
| | | | |
Db 1561 TGGGTGCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
| | | | |
QY 2143 ACACTGACTTCAAGACCAACCAACCCCTGCTCAATTAACCTTGGCCAGGCTCTCCAA 2202
| | | | |
Db 1621 ACACTGACTTCAAGACCAACCAACCCCTGCTCAATTAACCTTGGCCAGGCTCTCCAA 1680
| | | | |
QY 2203 GCTCAGCATTTGGGACCCAGCTCAGAGACCTGCTCTGCTGCTGCTGCTGCTGCTGCT 2262
| | | | |
Db 1681 GCTCAGCATTTGGGACCCAGCTCAGAGACCTGCTCTGCTGCTGCTGCTGCTGCTGCT 1740
| | | | |
QY 2263 CAAAAAAGACCTTTTCAATGATCAGATGACCTCAGAGAGATGGGTATGATGGGGCATC 2322
| | | | |
Db 1741 CAAAAAAGACCTTTTCAATGATCAGATGACCTCAGAGAGATGGGTATGATGGGGCATC 1800
| | | | |
QY 2323 ATCTCAGCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCCTGAGCTACAGC 2382
| | | | |
Db 1801 ATCTCAGCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCCTGAGCTACAGC 1860
| | | | |
QY 2383 TTCAATCACTCTGTTTCCAGAGATTTCTTTCAGCAATGCTATGCTTTCAGAGATGAG 2442
| | | | |
Db 1861 TTCAATCACTCTGTTTCCAGAGATTTCTTTCAGCAATGCTATGCTTTCAGAGATGAG 1920
| | | | |
QY 2443 AAGGGAGAGGTAAACATTTCTATTTGATATTTGAAAGAGCTGAAAGATAT 2502
| | | | |
Db 1921 AAGGGAGAGGTAAACATTTCTATTTGATATTTGAAAGAGCTGAAAGATAT 1980
| | | | |
QY 2503 GGAATACATGAGCTGTTTGGGGCATCAACACAGTTTCTATTTGGGCTGTTAAGTAT 2562
| | | | |
Db 1981 GGAATACATGAGCTGTTTGGGGCATCAACACAGTTTCTATTTGGGCTGTTAAGTAT 2040
| | | | |
QY 2563 GAGGGGAGAGAGATGAGAAACATCTTCACTGCGGCTGCTCAAGGGAGAGAACTG 2622
| | | | |
Db 2041 GAGGGGAGAGAGATGAGAAACATCTTCACTGCGGCTGCTCAAGGGAGAGAACTG 2100
| | | | |
QY 2623 ATGCAATGGGTCCCGTCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2682
| | | | |
Db 2101 ATGCAATGGGTCCCGTCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
| | | | |
QY 2683 TGCCTGACAGACTGAGAAACAAAAGCTTCTGACCAAGTGAATGGCCATTTGAGAA 2742
| | | | |
Db 2161 TGCCTGACAGACTGAGAAACAAAAGCTTCTGACCAAGTGAATGGCCATTTGAGAA 2220
| | | | |
QY 2743 ATGGGAGTGTGTAGAAACAGACATGAGCTTGTAGTGTGACATTTCTGACATTAATTC 2802
| | | | |
Db 2221 ATGGGAGTGTGTAGAAACAGACATGAGCTTGTAGTGTGACATTTCTGACATTAATTC 2280
| | | | |
QY 2803 AGCGGCACTGTAAAGAGCTTCAAGCTGATTTGAGGAGAGAGACAGACATTAAGTGAAGC 2862
| | | | |
Db 2281 AGCGGCACTGTAAAGAGCTTCAAGCTGATTTGAGGAGAGAGACAGACATTAAGTGAAGC 2340
| | | | |
QY 2863 CCCACATGATGCTGCTTCAAGTGTGAGGAGTCCAGTCAACAGATGCTTATTTGGAGATCTC 2922
| | | | |
Db 2341 CCCACATGATGCTGCTTCAAGTGTGAGGAGTCCAGTCAACAGATGCTTATTTGGAGATCTC 2400
| | | | |
QY 2923 TTCTCGTCTCAAGGTCAACAGAAACCTGAGAGCTGAGCCTTAAGTGAACCTGCTG 2982
| | | | |
Db 2401 TTCTCGTCTCAAGGTCAACAGAAACCTGAGAGCTGAGCCTTAAGTGAACCTGCTG 2460
| | | | |
QY 2983 AGGCACTCTGCAATGAAAGGCTTTGTGAAGACCTGAGAGGCTGCTGCTGCTGCTGCTG 3042
| | | | |
Db 2461 AGGCACTCTGCAATGAAAGGCTTTGTGAAGACCTGAGAGGCTGCTGCTGCTGCTGCTG 2520
| | | | |
QY 3043 ACCCTGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
| | | | |

[illegible][illegible]


```

/mo1_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/cell_type="activated Jurkat T cells"
1..4332
/gene="NAC"
1..4332
/gene="NAC"
/notes="Apaf-1/Ced-4 family apoptosis regulator: contains
NB-ARC domain, CARD domain, and leucine-rich repeats
(LRR)"
/codon_start=1
/product="NAC-delta splice variant"
/protein_id="AAK00751.1"
/db_xref="GI:12656111"
/translation="MAGAWGRRLACYLEFLKEELKEFOLLANKHSSSSGEPQAQ
PEKTSQMEVASIYAOYGEORAMDALHTMEOMGLRELCAOAGSHSPSPSPSE
PHLGSPOPTSTAVLMEIHLELPAQCTGSEERVLRLDPTSGRMEISASLYOL
PSSPDHSPSESPNAPTSTAVLGSWSPPOSLAPDEORAPGOWPLDSTGIVYE
IRRRERKSKGRPRMAAVVGTPOATSIQPHHNPBVSRESCSTWPKMKNPNO
KEFQULLOKPHPSQDPLVRSWPDVEENRHLIELRLFGGLDTPRVLVLGG
AAOIGKSTLARQVBEAWRGQLYGDPRQHFYFSCRELAOSKVALSLELIGKGTAP
APRLOISRPERLFLIDGVDEBQWVLOEBSSELCLMSQPOPADLGLSLGTLIP
EAEFLIARTTALONLPSLEOARWVEVLEFSESSREKEYFARYPTDEROAIAPRLVK
SNEIWAICLVPMWSWLAQCIQMOOKRKKLTLSTTTTCLHAYQAOLQAPLGP
QLEDLGSLAEGLWQKTLFSPDRLRKHLDGALISTFLMGLIOEHPIFLSTFTIL
CPQEFPAASVYLEDERKGRKHSNCLIDELKTEAYGHIHFGASTRFLGLSLDSE
EREMENIFHRLSGRNIMQWVPSILOLPHSLSESLHCLYETNKTFLQVNAHFE
MGKCVETMELVCTFCIFSRHYKQLLEGROHRSSTWSTPVLFRVYVDAVVO
ILFSYLVKTRNLKELDLGNSLSHSAVSKLCTLRRLRCLLETLRLAGLPTDCKD
LAGLRANOTLTELDSLSPVLTLAGAHLICORLPOPCUKORLOVSCGLTSCCOL
ASVLSASPSLKEIDLQONNLDIVGRLCSLRHPAKLRLGSPVMTTBSGLDTE
MSNSTKFORLGSERASHVAQANIKLDVSKIPIABIABESSSEVVEVELLQV
SPASQGLHRLPLGTDDEDFWGPPTGPAVEVDKKNYRHFPAAGSVRPNPGLCV
MRSAVYVEIFECVWDFLEGINPQHSWVAVGLDILKABGAVAVALPHFVALQSH
VDTSLFQMAHKEKEMLEKPARVELHIVLENSPSPSLGVALMINALRFLPTVS
VILYHRYHREVEHLYLIPSDCSIRKIDLEMKRPPVILHPRPLTPRYMGRYV
SGSGSMLEILPKRELECYRSPEDQLPSEFYVHLSGRLQYKDKEDTLWEALY
KPGDMLPATTLLIPPARIAVPSPLDAPQLHVRVDYRQQLIARVTSVVALDKLQVYL
SQDQYERVLAEHTRPSQMRKLFSLQSDMRCKDGLVALKETHPHILMEWKSXK
GLPLPS8"
BASE COUNT      1018 a      1202 c      1211 g      901 t
ORIGIN
Query Match      72.7%; Score 3958; DB 9; Length 4332;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 4200; Conservative 0; Mismatches 0; Indels 222; Gaps 2;
QY 523 ATGGCTGGCGGAGCTGGGGCGGCTGGCCCTGTTACTTGGAGTTCCTGAAGAAGAGAGAG 582
    |||
DB 1 ATGGCTGGCGGAGCTGGGGCGGCTGGCCCTGTTACTTGGAGTTCCTGAAGAAGAGAGAG 60
QY 583 CTGAAGAGAGTTCAGCTTCTGCTGCGCAATTAACGCACTCCAGAGACTCTTCCGGGTAG 642
    |||
DB 61 CTGAAGAGAGTTCAGCTTCTGCTGCGCAATTAACGCACTCCAGAGACTCTTCCGGGTAG 120
QY 643 ACACCCGCTCAGCCAGAGAAAGACAGTGGAGTGGCTGCTGTAACCTGGTGGCTCAG 702
    |||
DB 121 ACACCCGCTCAGCCAGAGAAAGACAGTGGAGTGGCTGCTGTAACCTGGTGGCTCAG 180
QY 703 TATGGGAGAGCAGCGGGGCTTGGGACCTAGCCCTCATACTCTGGAGAGCAATAGGGGCTAGG 762
    |||
DB 181 TATGGGAGAGCAGCGGGGCTTGGGACCTAGCCCTCATACTCTGGAGAGCAATAGGGGCTAGG 240
QY 763 TCACCTGTCGCGCCAGAGCCAGAGAAAGGGGAGGACACTCTCCCTATTCCCTTAGAGCCA 822
    |||
DB 241 TCACCTGTCGCGCCAGAGCCAGAGAAAGGGGAGGACACTCTCCCTATTCCCTTAGAGCCA 300
QY 823 AGTGAACCCCACTGGGGGTCTTCCAGCCCACTTCACCGGACGAGTGTATAGTCCCTGG 882
    |||
DB 301 AGTGAACCCCACTGGGGGTCTTCCAGCCCACTTCACCGGACGAGTGTATAGTCCCTGG 360
QY 883 ATCATGAAATTGGCGGGGGGTGCACCAGGAGCTCAAGAGAAAGGTTTGAACAAGTGT 942

```

```

DB 361 ATCATGAAATTGGCGGGGGGTGCACCAGGAGCTCAAGAGAAAGGTTTGAACAAGTGT 420
QY 943 CTGAACATCTGAGACGCGCGCTGGAAGAAATCTTGCTCACTCTCTTACCAAGCTCTT 1002
    |||
DB 421 CTGAACATCTGAGACGCGCGCTGGAAGAAATCTTGCTCACTCTCTTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCCAACCATTAGTGTCTTCAAGCCAGAGAGTCAACCCCAAGCCCCCACTCCACA 1082
    |||
DB 481 CCAAGCTCCCCAACCATTAGTGTCTTCAAGCCAGAGAGTCAACCCCAAGCCCCCACTCCACA 540
QY 1063 GCAGTGCTGGGAGAGCTGGGAGATCCCACTCAGGCCAGCTTAGACCCAGAGAGAGAG 1122
    |||
DB 541 GCAGTGCTGGGAGAGCTGGGAGATCCCACTCAGGCCAGCTTAGACCCAGAGAGAGAG 600
QY 1123 GCTCTTGGGACCCCAATGGCTCTTGATGAACGTCAGAAATTTTACTACAGAAATCAGA 1182
    |||
DB 601 GCTCTTGGGACCCCAATGGCTCTTGATGAACGTCAGAAATTTTACTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGACGCCCCCATGGGACGCGTGTGTAGAACG 1242
    |||
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGGACGCCCCCATGGGACGCGTGTGTAGAACG 720
QY 1243 CCCCCAAGGCGGACACACAGCTTACAGCCGCCACCAACCATGGGAGCCTTCTGTAGA 1302
    |||
DB 721 CCCCCAAGGCGGACACACAGCTTACAGCCGCCACCAACCATGGGAGCCTTCTGTAGA 780
QY 1303 GAGAGCTCTGTTTCCACATGGGCGCTGGAAAAATAGAGATTTTAAACAAAAATTCACAG 1362
    |||
DB 781 GAGAGCTCTGTTTCCACATGGGCGCTGGAAAAATAGAGATTTTAAACAAAAATTCACAG 840
QY 1363 CTGCTACTTTTCAAAAGACTCACCCCAAGAGCCAGATCCCTGGTCAAGAGAGCTGG 1442
    |||
DB 841 CTGCTACTTTTCAAAAGACTCACCCCAAGAGCCAGATCCCTGGTCAAGAGAGCTGG 900
QY 1423 CCGATATATGGAAGAGAAATCGAGAGCACTTATAGATTCAGAGACTTATTTGGGCCA 1482
    |||
DB 901 CCGATATATGGAAGAGAAATCGAGAGCACTTATAGATTCAGAGACTTATTTGGGCCA 960
QY 1483 GGCCTGTGATCCCAAGAACTTCGATAGTATACCTGAGGGGGCTGCTGGAATTTGGAG 1542
    |||
DB 961 GGCCTGTGATCCCAAGAACTTCGATAGTATACCTGAGGGGGCTGCTGGAATTTGGAG 1020
QY 1543 TCAACACTGGCCAGGACAGGTGGAAGAAAGCCTGGGGGAGAGAGCCAGCTGTATGGGAGACCG 1602
    |||
DB 1021 TCAACACTGGCCAGGACAGGTGGAAGAAAGCCTGGGGGAGAGAGCCAGCTGTATGGGAGACCG 1080
QY 1603 TTCCAGCATGCTTTCTTCACTTCAAGTGCAGAGAGTGGCCAGTCCAGAGTGTAGTCTC 1662
    |||
DB 1081 TTCCAGCATGCTTTCTTCACTTCAAGTGCAGAGAGTGGCCAGTCCAGAGTGTAGTCTC 1140
QY 1663 GCTGAGCTCATTCGAAAAAGATGGGACAGCCACTTCGGCTCCATTAGACAGATCTGTCT 1722
    |||
DB 1141 GCTGAGCTCATTCGAAAAAGATGGGACAGCCACTTCGGCTCCATTAGACAGATCTGTCT 1200
QY 1723 AGGCAGAGAGCGCTGCTCTTCACTCTGATGCTGTATGATGAGCCAGATGGGTCTTGACAG 1782
    |||
DB 1201 AGGCAGAGAGCGCTGCTCTTCACTCTGATGCTGTATGATGAGCCAGATGGGTCTTGACAG 1260
QY 1783 GAGCGGAGTCTGAGAGCTGCTGTCAGACTGAGAGCCAGAGCCAGCGGGGAGATGCACTGCTG 1842
    |||
DB 1261 GAGCGGAGTCTGAGAGCTGCTGTCAGACTGAGAGCCAGAGCCAGCGGGGAGATGCACTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAACTATATCTTCCGAGGAGATCTTCTGATCAGGCTCGGAGCC 1902
    |||
DB 1321 GGCAGTTTGTGGGAAAACTATATCTTCCGAGGAGATCTTCTGATCAGGCTCGGAGCC 1380
QY 1903 ACAAGCTTGAAGAACTCATTTCTTCTTGGAGCAGGCAAGTGTGGATGAGTCTTGGG 1962
    |||
DB 1381 ACAAGCTTGAAGAACTCATTTCTTCTTGGAGCAGGCAAGTGTGGATGAGTCTTGGG 1440
QY 1963 TTCTGTAGTCCAGAGAGAAATATTTTCAAGATATTTTCAAGATGGAAGGCAAGCA 2022
    |||

```

Db 1441 TTCTGAGTCCAGCAGAGAAATATTTCTACAGATATTTCAAGATGAAAGCAGCA 1500
 QY 2023 ATTAGACCTTTAGTTGGTGTCAATATCAAAAGAGCTCTGAGCCCTGTGTCTTGTC 2082
 Db 1501 ATTAGAGCCCTTTAGTTGGTGTCAATATCAAAAGAGCTCTGAGCCCTGTGTCTTGTC 1560
 QY 2083 TGGGTGCTGAGTGGCTGAGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2142
 Db 1561 TGGGTGCTGAGTGGCTGAGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1620
 QY 2143 ACATGACTTCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2202
 Db 1621 ACATGACTTCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1680
 QY 2203 GCTGAGCCTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2262
 Db 1681 GCTGAGCCTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1740
 QY 2263 CAAAAAAGACCTTTTCAAGTCCAGATGAGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322
 Db 1741 CAAAAAAGACCTTTTCAAGTCCAGATGAGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 QY 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCAGCAGCAGCAGCAGCAGCAGC 2382
 Db 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCAGCAGCAGCAGCAGCAGCAGC 1860
 QY 2383 TTCTATTCACCTCTGTTTCCAGAGCTTTTGAAGATGGGTATTTCTTCAAGAGATGAG 2442
 Db 1861 TTCTATTCACCTCTGTTTCCAGAGCTTTTGAAGATGGGTATTTCTTCAAGAGATGAG 1920
 QY 2443 AAGGGAGAGATTAATTTATTTGATGATGATTTGAAAAGAGAGAGAGAGAGAGAGATTT 2502
 Db 1921 AAGGGAGAGATTAATTTATTTGATGATGATTTGAAAAGAGAGAGAGAGAGAGATTT 1980
 QY 2503 GGAATACATGCTGTTTGGGAGATCAACACAGCTTTCTATTTGGGCTGTATGATGAT 2562
 Db 1981 GGAATACATGCTGTTTGGGAGATCAACACAGCTTTCTATTTGGGCTGTATGATGAT 2040
 QY 2563 GAGGGAGAGAGAGATGAGAAATCTTTCACTCCGCTGTCTCAAGGGAGAGAGAGAGAG 2622
 Db 2041 GAGGGAGAGAGAGATGAGAAATCTTTCACTCCGCTGTCTCAAGGGAGAGAGAGAGAG 2100
 QY 2623 ATGACAGTGGGTCCGCTCCCTGAGCTGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2682
 Db 2101 ATGACAGTGGGTCCGCTCCCTGAGCTGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
 QY 2683 TGCCTTGACAGACTCGGACCAAAAGCTTCTGACACAGTGAATGGCCATTTGAGAGAA 2742
 Db 2161 TGCCTTGACAGACTCGGACCAAAAGCTTCTGACACAGTGAATGGCCATTTGAGAGAA 2220
 QY 2743 ATGGGACATGTTGTAAGAACAGACATGAGAGCTTTAGTGTGCACTTTTGATTAATTC 2802
 Db 2221 ATGGGACATGTTGTAAGAACAGACATGAGAGCTTTAGTGTGCACTTTTGATTAATTC 2280
 QY 2803 AGCGGACGCTGAAGAGCTTCACTGATTTGAGGGCAGGAGAGAGAGAGAGAGAGAGAGAG 2862
 Db 2281 AGCGGACGCTGAAGAGCTTCACTGATTTGAGGGCAGGAGAGAGAGAGAGAGAGAGAGAG 2340
 QY 2863 CCCACCATGATGATCTGTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGAGATTTCTC 2922
 Db 2341 CCCACCATGATGATCTGTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGAGATTTCTC 2400
 QY 2923 TTCTCCGCTCTCAAGATCAGCAGAACTGAAGAGAGCTGAAGCTTAAGTGAAGAGAGAG 2982
 Db 2401 TTCTCCGCTCTCAAGATCAGCAGAACTGAAGAGAGCTGAAGCTTAAGTGAAGAGAGAG 2460
 QY 2983 AGCCACCTGTCAGAGAGAGATCTTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3042
 Db 2461 AGCCACCTGTCAGAGAGAGATCTTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
 QY 3043 ACCCTGGAGTTGGCTGGCTGGCTGACAGCTGAGAGAGCTGAAGAGAGAGAGAGAGAG 3102
 Db 2521 ACCCTGGAGTTGGCTGGCTGGCTGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580

QY 3103 CTGAGAGCCACCAAGCCTGACCGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 3162
 Db 2581 CTGAGAGCCACCAAGCCTGACCGAGCTGAGCTGAGCTTCAATGTGCTCAAGATGCT 2640
 QY 3163 GAGAGCCAAACACCTTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3222
 Db 2641 GAGAGCCAAACACCTTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
 QY 3223 CTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3282
 Db 2701 CTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2760
 QY 3283 AGCCCAAGCTGAAG 3342
 Db 2761 AGCCCAAGCTGAAG 2820
 QY 3343 CTGCTGAG 3402
 Db 2821 CTGCTGAG 2869
 QY 3403 ACAACTCTGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3462
 Db 2870 -----GGAACCAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2869
 QY 3463 CTGCTCATTTTCAAG 3522
 Db 2870 -----GGAACCAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2910
 QY 3523 GAGAGAGATGATTAATGAG 3582
 Db 2911 GAGAGAGATGATTAATGAG 2970
 QY 3583 GCTTCCATGTTGCTCAGAGCTATCTCAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 3642
 Db 2971 GCTTCCATGTTGCTCAGAGCTATCTCAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 3030
 QY 3643 GCTGAGATTTGAG 3702
 Db 3031 GCTGAGATTTGAG 3090
 QY 3703 TCTCCGCTCTCAAG 3762
 Db 3091 TCTCCGCTCTCAAG 3150
 QY 3763 GAGCCCAAG 3822
 Db 3151 GAGCCCAAG 3210
 QY 3823 CACTTCCCTGAG 3882
 Db 3211 CACTTCCCTGAG 3270
 QY 3883 GAAAGCGATGAG 3942
 Db 3271 GAAAGCGATGAG 3330
 QY 3943 CCAAGACAGAGTGAATGAG 4002
 Db 3331 CCAAGACAGAGTGAATGAG 3390
 QY 4003 GTGAGAGCTGTGACCTTCCCTCACTTTGTGCTTCCAAAGGGGGCCATGTGAGACATATC 4062
 Db 3391 GTGAGAGCTGTGACCTTCCCTCACTTTGTGCTTCCAAAGGGGGCCATGTGAGACATATC 3450
 QY 4063 CTGTTTCAAAATGAG 4122
 Db 3451 CTGTTTCAAAATGAG 3510
 QY 4123 GAGCTGATCAGATGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4182
 Db 3511 GAGCTGATCAGATGATGTTGAG 3570


```

Db      481 CCAGGCTCCCAAGACCATGAGTCTCCAGCGAGAGTCACCCAAGCCCCCAATCCACA 540
Qy      1063 GCAGTCTGGGAGAGCTGGGGATCCCACTCAAGCCAGCCTTAGCACCCGAGAGAGAG 1122
Db      541 GAGTGTCTGGGAGCTGGGGATCCCACTCAAGCCAGCCTTAGCACCCGAGAGAGAG 600
Qy      1123 GCTCCTGGGAGCCCAATGGGCTCTGATGAAAGCTCAGAAATTTACTACAGAAATCAGA 1182
Db      601 GCTCTGGGAGCCCAATGGGCTCTGATGAAAGCTCAGAAATTTACTACAGAAATCAGA 660
Qy      1183 GAAAGAGAGAGAGAAATCAGAGAAAGGCAAGGCCCAATGGGCAAGCGTGTAGAGAG 1242
Db      661 GAAAGAGAGAGAGAAATCAGAGAAAGGCAAGGCCCAATGGGCAAGCGTGTAGAGAG 720
Qy      1243 CCCCCCAGGCGGCAACCAAGCCTTACAGCCCAACCAACCCATGGAGGCTTCTGTAGA 1302
Db      721 CCCCCCAGGCGGCAACCAAGCCTTACAGCCCAACCAACCCATGGAGGCTTCTGTAGA 780
Qy      1303 GAGAGCCTCTGTTCACATGAGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACAG 1362
Db      781 GAGAGCCTCTGTTCACATGAGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACAG 840
Qy      1363 CTGCTACTTCTACAAAGACCTCAACCCAGAAAGCCAAAGATCCCTGTCTAAGAGAGCTG 1422
Db      841 CTGCTACTTCTACAAAGACCTCAACCCAGAAAGCCAAAGATCCCTGTCTAAGAGAGCTG 900
Qy      1423 CCTGATTAATGTAGAGAGAGATCAGAGACATTTAATTTAGATCAGAGACTTTATGGCCA 1482
Db      901 CCTGATTAATGTAGAGAGAGATCAGAGACATTTAATTTAGATCAGAGACTTTATGGCCA 960
Qy      1483 GGCCTGGATACCCAAAGACCTCGCATAGTCACTGACAGGAGGCTGTGAAATTGGGAG 1542
Db      961 GGCCTGGATACCCAAAGACCTCGCATAGTCACTGACAGGAGGCTGTGAAATTGGGAG 1020
Qy      1543 TCAACACTGGCCAGGAGGTGAAGAAAGCTGGGGAGAGAGCCAGCTGTATGGGAGCCG 1602
Db      1021 TCAACACTGGCCAGGAGGTGAAGAAAGCTGGGGAGAGAGCCAGCTGTATGGGAGCCG 1080
Qy      1603 TTCCAGACATGTCTTCTTACTTCAAGTCAAGAGCTGAGAGCTGCCAGTCAAGGTGTGAGTCTC 1662
Db      1081 TTCCAGACATGTCTTCTTACTTCAAGTCAAGAGCTGAGAGCTGCCAGTCAAGGTGTGAGTCTC 1140
Qy      1663 GCTGAGCTCATCGGAAAAAGATGGGACAGGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
Db      1141 GCTGAGCTCATCGGAAAAAGATGGGACAGGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
Qy      1723 AGGCGAAGCGGCTGTCTTCTTCACTGATGTGTAGATGAGCCAGATGGTCTTTGCA 1782
Db      1201 AGGCGAAGCGGCTGTCTTCTTCACTGATGTGTAGATGAGCCAGATGGTCTTTGCA 1260
Qy      1783 GAGCCGATTTCTGAGCTCTGTCTGCACTGAGCCAGCCACAGCGCGCGGATGCACTGCTG 1842
Db      1261 GAGCCGATTTCTGAGCTCTGTCTGCACTGAGCCAGCCACAGCGCGCGGATGCACTGCTG 1320
Qy      1843 GGCAGTTTCTGGGAAAAACTATATCTCCGAGGACCTTCTCTGATCAAGCTCGGAGCC 1902
Db      1321 GGCAGTTTCTGGGAAAAACTATATCTCCGAGGACCTTCTCTGATCAAGCTCGGAGCC 1380
Qy      1903 ACAGCTCTGAGAACTCATTTCTTTTGGAGCAGGACGTTGGGTAGAGGTCTGGGG 1962
Db      1381 ACAGCTCTGAGAACTCATTTCTTTTGGAGCAGGACGTTGGGTAGAGGTCTGGGG 1440
Qy      1963 TTCTCTGAGTCCGAGAGAGAAATTTTCTACAGATTTTCACAATGAAAGGAGAGA 2022
Db      1441 TTCTCTGAGTCCGAGAGAGAAATTTTCTACAGATTTTCACAATGAAAGGAGAGA 1500
Qy      2023 ATTAGAGCCTTTAGTGTGCTCAATCAACAAAGAGCTGTGGCCTGTGTCTTGTGCC 2082
Db      1501 ATTAGAGCCTTTAGTGTGCTCAATCAACAAAGAGCTGTGGCCTGTGTCTTGTGCC 1560
Qy      2083 TGGGTCTCTGGCTGGCTGCACTTGTCTGATCAGAGATGAAGCGGAGGAAAACTC 2142
Db      1561 TGGGTCTCTGGCTGGCTGCACTTGTCTGATCAGAGATGAAGCGGAGGAAAACTC 1620

```

```

Qy      2143 AACATGACTTCCAAAGACCAACCAACCTCTGTCTTACTTACCTTGGCCCAAGCTCTCCAA 2202
Db      1621 AACATGACTTCCAAAGACCAACCAACCTCTGTCTTACTTACCTTGGCCCAAGCTCTCCAA 1680
Qy      2203 GCTCAGCATTGGGAGCCCGAGCTCAGAGACCTCTCTCTGAGCTGTGAGGAGCATCTGG 2262
Db      1681 GCTCAGCATTGGGAGCCCGAGCTCAGAGACCTCTCTCTGAGCTGTGAGGAGCATCTGG 1740
Qy      2263 CAAAAAAGACCTTTTCAATGCTCAGATGACCTCAGAAAGCATGGGTATGATGGGCCATC 2322
Db      1741 CAAAAAAGACCTTTTCAATGCTCAGATGACCTCAGAAAGCATGGGTATGATGGGCCATC 1800
Qy      2323 ATCTCCACTTCTTGAATAGGTATTTCTTCAAGAGCAACCCATGCCCTGAGCTACAGC 2382
Db      1801 ATCTCCACTTCTTGAATAGGTATTTCTTCAAGAGCAACCCATGCCCTGAGCTACAGC 1860
Qy      2383 TTCATTCACTCTGTCTTCAAGAGTCTTTCAGAGCAAGTCCATATGTCTTGGAGATGAG 2442
Db      1861 TTCATTCACTCTGTCTTCAAGAGTCTTTCAGAGCAAGTCCATATGTCTTGGAGATGAG 1920
Qy      2443 AAGGAGAGGTAAACATTTATTTGATCATATGATTTGAAAGACGCTAGAGCATAT 2502
Db      1921 AAGGAGAGGTAAACATTTATTTGATCATATGATTTGAAAGACGCTAGAGCATAT 1980
Qy      2503 GGAATACATGGCTGGTGGGCAATCAACACGTTTCCATTTGGGCTGTATGAT 2562
Db      1981 GGAATACATGGCTGGTGGGCAATCAACACGTTTCCATTTGGGCTGTATGAT 2040
Qy      2563 GAGGGGAGAGAGATGAGAGACATCTTCACTGCGGCTGTCTCAGGGGAGAACTG 2622
Db      2041 GAGGGGAGAGAGATGAGAGACATCTTCACTGCGGCTGTCTCAGGGGAGAACTG 2100
Qy      2623 ATGCAATGGTCCGCTCCCTGAGCTGTCTGTGAGCCACATCTCTGTGAGTCCCTCAC 2682
Db      2101 ATGCAATGGTCCGCTCCCTGAGCTGTCTGTGAGCCACATCTCTGTGAGTCCCTCAC 2160
Qy      2683 TGCTTGTAAGAGACTCGGAACAAACGTTCCGTAACAAGATGAGGCCATTTGAGAA 2742
Db      2161 TGCTTGTAAGAGACTCGGAACAAACGTTCCGTAACAAGATGAGGCCATTTGAGAA 2220
Qy      2743 ATGGCATGTGTGTAGAAACAGACATGAGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2802
Db      2221 ATGGCATGTGTGTAGAAACAGACATGAGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2280
Qy      2803 AGCCGCCACGTGAGAGAGCTTCACTGATTGAGGCCAGGCAAGATCAACATGAGAGC 2862
Db      2281 AGCCGCCACGTGAGAGAGCTTCACTGATTGAGGCCAGGCAAGATCAACATGAGAGC 2340
Qy      2863 CCCACATGTGTGTGTGAGTGTGAGGAGTCCAGTCAAGATGAGTGTGAGATTC 2922
Db      2341 CCCACATGTGTGTGTGAGTGTGAGGAGTCCAGTCAAGATGAGTGTGAGATTC 2400
Qy      2923 TTCTCCGTCTTCAAGTCAACGAAACCTGAAAGAGCTGGAACCTTAATGAGAACTCGCTG 2982
Db      2401 TTCTCCGTCTTCAAGTCAACGAAACCTGAAAGAGCTGGAACCTTAATGAGAACTCGCTG 2460
Qy      2983 AGCCACTCTGCACTGAAAGTCTTTGTAGAACCTTGAAGACGCTCTGCTCTCTGAG 3042
Db      2461 AGCCACTCTGCACTGAAAGTCTTTGTAGAACCTTGAAGACGCTCTGCTCTCTGAG 2520
Qy      3043 ACCCTGGGTTGGCTGGGCTGAGGCTCAAGCTGAGAGCTGCAAGAGACCTTGCCTTTGGG 3102
Db      2521 ACCCTGGGTTGGCTGGGCTGAGGCTCAAGCTGAGAGCTGCAAGAGACCTTGCCTTTGGG 2580
Qy      3103 CTGAGAGCCAAACAGACCTCTGACGAGCTGGAACCTTGAATGTGCTACGAGATGT 3162
Db      2581 CTGAGAGCCAAACAGACCTCTGACGAGCTGGAACCTTGAATGTGCTACGAGATGT 2640
Qy      3163 GAGGCGAAACACCTTTGCAAGAGATGAGACAGCGAGCTGGAACCTTCAAGGACCTGAG 3222
Db      2641 GAGGCGAAACACCTTTGCAAGAGATGAGACAGCGAGCTGGAACCTTCAAGGACCTGAG 2700

```

QY 3223 CTGCTCAGCTGAGCCCTCACGTCGACCTGTCGACGAGCACTGGGCTCTGCTTATGAGCC 3282
 DB 2701 CTGTCAGCTGTGAGCCCTCACGTCGACCTGTCGACGAGCACTGGGCTCTGCTTATGAGCC 2760
 QY 3283 AGCCCAAGCTGAGAGGAGTGAAGCTGTCGACAGAAACAACCTGATGACCTTGGCGTGGCA 3342
 DB 2761 AGCCCAAGCTGAGAGGAGTGAAGCTGTCGACAGAAACAACCTGATGACCTTGGCGTGGCA 2820
 QY 3343 CTGCTCTGAGAGGAGCTCAGGCACTCTGCTGCAAACTCAATGAGCTGGGGCTGGACCAAG 3402
 DB 2821 CTGCTCTGAGAGGAGCTCAGGCACTCTGCTGCAAACTCAATGAGCTGGGGCTGGACCAAG 2867
 QY 3403 ACAACTGTAGTATGATGAGTGAAGAGCACTGAGGGGCTTGAAGAGAGAAACCTCAG 3462
 DB 2868 ----- 2867
 QY 3463 CTGCTCATCTTCAGCAGACCGAAACAAGTGTATGACCCCTTACTGAGGGCTGTGATACG 3522
 DB 2868 -----GAAACCAAGTGTATGACCCCTTACTGAGGGCTGTGATACG 2907
 QY 3523 GGAAGATGATATATGACATCTCTCACTCAAGCGGAGAGACTCGATCAGAGAGGGCG 3582
 DB 2908 GGAAGATGATATATGACATCTCTCACTCAAGCGGAGAGACTCGATCAGAGAGGGCG 2967
 QY 3583 GCTTCCCATGTTGCTCAGGCTATCTCAACTCTCTGAGCGTGAAGATCTTCCCAATT 3642
 DB 2968 GCTTCCCATGTTGCTCAGGCTATCTCAACTCTCTGAGCGTGAAGATCTTCCCAATT 3027
 QY 3643 GCTGAGATTTGACAGAGAAAGCTCCCGAGAGTATACCGGTGAACTTTGCTGCTGCT 3702
 DB 3028 GCTGAGATTTGACAGAGAAAGCTCCCGAGAGTATACCGGTGAACTTTGCTGCTGCTGCT 3087
 QY 3703 TCTCTGCTCTCTCAAGGGAGCTTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGG 3762
 DB 3088 TCTCTGCTCTCTCAAGGGAGCTTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGG 3147
 QY 3763 GGGCCCAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3822
 DB 3148 GGGCCCAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3207
 QY 3823 CACTTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3882
 DB 3208 CACTTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3267
 QY 3883 GAAGCGGTGACCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3942
 DB 3268 GAAGCGGTGACCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3327
 QY 3943 CCAACAGACAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4002
 DB 3328 CCAACAGACAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3387
 QY 4003 GTGGAAGCTGTGACCTCTCACTTGTGCTCTCTCAAGGGGGCCATGTGACACATCC 4062
 DB 3388 GTGGAAGCTGTGACCTCTCACTTGTGCTCTCTCAAGGGGGCCATGTGACACATCC 3447
 QY 4063 CTGTTCCAAATGAGCCCACTTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCCAGCGAGGTG 4122
 DB 3448 CTGTTCCAAATGAGCCCACTTTAAAGAGAGAGGAGTGTCTCTGAGAAAGCCAGCGAGGTG 3507
 QY 4123 GAGCTGCATACATGATGTTCTGAGAAACCCAGCTTCTCCCTTGGAGAGTCCCTGAGAA 4182
 DB 3508 GAGCTGCATACATGATGTTCTGAGAAACCCAGCTTCTCCCTTGGAGAGTCCCTGAGAA 3567
 QY 4183 ATGATCATATATGCGCTGCGCTTCACTCCGTCACCTCTGTGTGTGCTTTACACCGC 4242
 DB 3568 ATGATCATATATGCGCTGCGCTTCACTCCGTCACCTCTGTGTGTGCTTTACACCGC 3627
 QY 4243 GTCCATCTTGAAGAGTGCACCTTCCACTTCACTGATCCCAAGTGAAGTGTCTCCATTCGG 4302
 DB 3628 GTCCATCTTGAAGAGTGCACCTTCCACTTCACTGATCCCAAGTGAAGTGTCTCCATTCGG 3687
 QY 4303 ----- 4302

DB 3688 AAGGCAATGATGATCTAATAATGAATTCAGTTTGTGCAATCCACAGCAACCCCGC 3747
 QY 4303 ----- 4302
 DB 3748 CTGACCCCACTTATATATGAGCTGTGTGTAACCTGTGTGGGTCTGGTTCAAGGATGCTG 3807
 QY 4303 -----AAGAACTGAGCTCTGTGTATGAAAGCCCTTGAAGAAAGCAAGCTGTC 4350
 DB 3808 GAAATATCTCCCAAGAACTGTGAGCTCTGTATGAAAGCCCTTGAAGAAAGCAAGCTGTC 3867
 QY 4351 TCGAGTTCTACGTTGAGCCCACTTGGATCAGGGATCAGGGCTGCAAGTGAAGAAAGCAAGAA 4410
 DB 3868 TCGAGTTCTACGTTGAGCCCACTTGGATCAGGGATCAGGGCTGCAAGTGAAGAAAGCAAGAA 3927
 QY 4411 GATGAGCTCTGTGTGTGAGAGGCTTGTGTAACCAAGAGATCTATGCTGCAACTACT 4470
 DB 3928 GATGAGCTCTGTGTGTGAGAGGCTTGTGTAACCAAGAGATCTATGCTGCAACTACT 3987
 QY 4471 CTGATCTCTCAGCCCGCATAGCCGTAACCTTCACTCTGTGATGCTCCCGCAGTTGCTGAC 4530
 DB 3988 CTGATCTCTCAGCCCGCATAGCCGTAACCTTCACTCTGTGATGCTCCCGCAGTTGCTGAC 4047
 QY 4531 TTTGTGACCAAGTATGAGAGAGCTGATAGCCCGATGATCGGGTGAAGTGTGCTTG 4590
 DB 4048 TTTGTGACCAAGTATGAGAGAGCTGATAGCCCGATGATCGGGTGAAGTGTGCTTG 4107
 QY 4591 GACAACTGATGACAGAGTGTGACAGCAGAGCACTAGAGAGGAGTGTGCTGAGAAC 4650
 DB 4108 GACAACTGATGACAGAGTGTGACAGCAGAGCACTAGAGAGGAGTGTGCTGAGAAC 4167
 QY 4651 ACAGAGCCCAAGCCAGATGCGGAAAGCTTGTGATGAGCACTCTGAGAGCCGAAAGTGC 4710
 DB 4168 ACAGAGCCCAAGCCAGATGCGGAAAGCTTGTGATGAGCACTCTGAGAGCCGAAAGTGC 4227
 QY 4711 AAGATGAGCTCTACCAAGCCCTGAGAGAGACCATCTCACTTATGAACTCTG 4770
 DB 4228 AAGATGAGCTCTACCAAGCCCTGAGAGAGACCATCTCACTTATGAACTCTG 4287
 QY 4771 GAGAAAGGAGCAAAAAAGGAGCTCTGCACTCAGCAGCTGA 4812
 DB 4288 GAGAAAGGAGCAAAAAAGGAGCTCTGCACTCAGCAGCTGA 4329

RESULT 14
 HSM800983
 LOCUS
 DEFINITION
 Homo sapiens mRNA; cDNA DKFZp58601822 (from clone DKFZp58601822); partial cds.
 ACCESSION
 AL117470
 VERSION
 AL117470.1 GI:5911938
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 4123)
 Koehler, K., Beyer, A., Mewes, H. W., Gassenhuber, J. and Wiemann, S.
 Direct Submission
 Submitted (15-SEP-1999) MIPS, Am Klopferpitze 18a, D-82152 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charité, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp58601822) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de
 Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 Location/Qualifiers
 1. .4123

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ58601822"
/issue_type="uteras"
/clone_lib="586 (synonym: hute1). Vector pSport1, host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1..4123
/gene="DKFZ58601822"
<1..3580
/gene="DKFZ58601822"
/feature="XIA0926, partly, alternative spliced"
/codon_start=-2
/product="hypothetical protein"
/protein_id="CAB55945.1"
/db_xref="GI:5911939"
/translation="LLLRPHRSDPLVKSMPDVEENRGLIIRDLFPGGLDLO
EBRIVLQAGIKSTLAROYKEMRGOLYGRROHAFPSCEBLQSKVSLAEI
IKDGTAPAPRIQLISREPLIFLDVDEBQWVLOPSSSLCHMSOPADALIG
SLGKTLPEASFLITARTALQNLIPSLQARWEVIGFSSSSKEYFYRYFTBERQ
ALRAFRLVSKNELVALCLVPWVSLACIQLQMKREKLTLSKTTTLCIHYLAQ
ALQAPLGPOLDLCSLAEGIMOKTIFSPDLRGHGDGAIISTFLMGILQHPPI
PLSYSPILHCFQEFPAASGYLDEKRGKHSNCIDLEKTLKAYGHGIFGASTTRF
LIGLSDSEGEREMENFICRLSQGRNLQMWPSLQILQPSLSEGLHCYETRNPTL
TVYMAHFEEMKCVETDMLICTPCTIKSRVYKQLQLEGRQHSSTSPVIVLFRW
VPTAYWQILPSVLKVTNLKELDLSGSLSHSAVKSICKLRLPRCLLETURLAGC
GLIADCKDLAFGLRANQTLLELDLSFNLTDGAKHLICRLRPSCKLQRLQVSCG
LTSDDCCDLASVLSASPSLKEHLDQNNLDVGRLLCGLHHPACKLRLDQDTLL
SDMEQNLALKEOKPOLIFSRKPSVMTPEGLDGTMSNSTSLKRGSGSRAA
SHVAQNLKLDLBSKIFPIAETAESSEPVVELLCVPSPASODLHTRKPLGDDF
WPTPTPVATEVVDKERNLYRYVPAVAGSRMPTGCTFMREAVTYEIEFCWDPLG
EINPQSMWVAPLBDIKAEPAVEAVHLPHFVALQGGVLDLQVAFKEEMILE
KARVELHNI VLENPSFPLGVLKMINHARFIPTVSVDLYHRHVEPFIHLYLI
PDDGIRKAIIDLKEMKFCVRIHKPEPLPLMYGCRVYVSGSGMLBILPELICY
RPPGDOLSEFVYGHLSGIRLOVYDKDDELTYMEALYKPDMLPATLLIPALIAV
PPLAPOLHFDVDOYREQLIARVTSVEVLKLGQVLSQSOBYRVALENRBQMR
KLPSLSQSDRCKRQGLYQALKEITHPHLIMELMEKSKKGLPLSS"
4057..4062
pol1A_signal
/gene="DKFZ58601822"
4081
pol1A_site
/gene="DKFZ58601822"
BASE COUNT 1002 a 1086 c 1136 g 899 t
ORIGIN
Query Match 69.9%; Score 3807.8; DB 9; Length 4123;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 3951; Conservative 0; Mismatches 2; Indels 132; Gaps 1;
QY 1365 GCTACTTTCACCAAGCCTTCAACCCAGAACCCAGATCCCTGTGTCAAGAGAGCTGGCC 1424
DB 1 GCTACTTTCACCAAGCCTTCAACCCAGAACCCAGATCCCTGTGTCAAGAGAGCTGGCC 60
QY 1425 TGATTATGTGAGAGAGATCGAGAGACATTTAATTGATCAGAGACTTATTTGGCCACAG 1484
DB 61 TGATTATGTGAGAGAGATCGAGAGACATTTAATTGATCAGAGACTTATTTGGCCACAG 120
QY 1485 CCTGATACCAAGAACCTTCAGTACTATCTGACAGGGGCTGCTGGAATTGGAGAGTC 1544
DB 121 CCTGATACCAAGAACCTTCAGTACTATCTGACAGGGGCTGCTGGAATTGGAGAGTC 180
QY 1545 AACACTGGCCAGGAGAGTGAAGAGACCTGGGGGAGAGGACCACTGTATAGGGGACGGCTT 1604
DB 181 AACACTGGCCAGGAGAGTGAAGAGACCTGGGGGAGAGGACCACTGTATAGGGGACGGCTT 240
QY 1605 CCAGCATGCTTTCTTACTTCACTGACGAGAGAGCTGGCCAGTCCCAAGGTGATGAGTCTCCG 1664
DB 241 CCAGCATGCTTTCTTACTTCACTGACGAGAGAGCTGGCCAGTCCCAAGGTGATGAGTCTCCG 300
QY 1665 TGAGCTCATCGAGAGAGATGGAGACAGCCACTCCGCTCCCATTTAGA CAAGTCTGTCTAG 1724
DB 301 TGAGCTCATCGAGAGAGATGGAGACAGCCACTCCGCTCCCATTTAGA CAAGTCTGTCTAG 360
```

```
QY 1725 GCCAGAGCGGCTGCTCTTCACTCTCGATGCTGTATAGAGCCAGATGGCTTCTGACAGA 1784
DB 361 GCCAGAGCGGCTGCTCTTCACTCTCGATGCTGTATAGAGCCAGATGGCTTCTGACAGA 420
QY 1785 GCCAGATTCAGAGCTCTGCTGCACTGAGAGCCAGGACAGCCGGGAGATGACTCTGGG 1844
DB 421 GCCAGATTCAGAGCTCTGCTGCACTGAGAGCCAGGACAGCCGGGAGATGACTCTGGG 480
QY 1845 CAGTTTGTGGGAGAAACTATATCTTCCGAGGACATCTTCTGTATACGGCTCGAGCAAC 1904
DB 481 CAGTTTGTGGGAGAAACTATATCTTCCGAGGACATCTTCTGTATACGGCTCGAGCAAC 540
QY 1905 AGCTTCGAGAGAACCTTATCTCTTGTGAGAGAGGACAGTTGGGTAGAGTCTGGGGT 1964
DB 541 AGCTTCGAGAGAACCTTATCTCTTGTGAGAGAGGACAGTTGGGTAGAGTCTGGGGT 600
QY 1965 CTCTGAGTCCAGCAGAGAGAAATATTTTCAAGATATTTTCAAGATGAGAGAGCAAGCAAT 2024
DB 601 CTCTGAGTCCAGCAGAGAGAAATATTTTCAAGATATTTTCAAGATGAGAGAGCAAGCAAT 660
QY 2025 TAGAGCTTTAGTGTGTCAAATCAACAAAGAGCTGGGCTGTGTCTTGTGCTCTG 2084
DB 661 TAGAGCTTTAGTGTGTCAAATCAACAAAGAGCTGGGCTGTGTCTTGTGCTCTG 720
QY 2085 GGTTGCTGGCTGGCTGCTGCTGCTGATGAGAGATGAGAGCGGAGAGAGAAATCTAC 2144
DB 721 GGTTGCTGGCTGGCTGCTGCTGCTGATGAGAGATGAGAGCGGAGAGAGAAATCTAC 780
QY 2145 ACTGACTTCCAGAGACCAACCAACCCCTGTGTCTATCATTAACCTTGGCCAGGCTCTCAAGC 2204
DB 781 ACTGACTTCCAGAGACCAACCAACCCCTGTGTCTATCATTAACCTTGGCCAGGCTCTCAAGC 840
QY 2205 TCAGCATTTGGAGACCCAGCTCAGAGACCTTGTCTGTGTGCTGAGGAGCATTTGGCA 2264
DB 841 TCAGCATTTGGAGACCCAGCTCAGAGACCTTGTCTGTGTGCTGAGGAGCATTTGGCA 900
QY 2265 AAAAAAGACCTTTTCACTCAGATGAGACCTCAGAGAGATGGGTAGATGGGGCCATCAT 2324
DB 901 AAAAAAGACCTTTTCACTCAGATGAGACCTCAGAGAGATGGGTAGATGGGGCCATCAT 960
QY 2325 CTCACACTTCTTGAAGATGGGTATCTTCAAGAGACCCCATCCCTGTAGAGCTACAGCTT 2384
DB 961 CTCACACTTCTTGAAGATGGGTATCTTCAAGAGACCCCATCCCTGTAGAGCTACAGCTT 1020
QY 2385 CATTCACCTCTGTCTTCAAGAGCTTCTTGCAGCAATGTCTTATGTCTTGGAGATGAGAA 2444
DB 1021 CATTCACCTCTGTCTTCAAGAGCTTCTTGCAGCAATGTCTTATGTCTTGGAGATGAGAA 1080
QY 2445 GGGGAGAGGTAAACATTCATATTGATGATGATTTGGAAAAAGAGCTTAGAGCATATGG 2504
DB 1081 GGGGAGAGGTAAACATTCATATTGATGATGATTTGGAAAAAGAGCTTAGAGCATATGG 1140
QY 2505 AATACATGGCTGTTTGGGGGATCAACACAGTTTCTATTTGGGCTGTATTAAGATGA 2564
DB 1141 AATACATGGCTGTTTGGGGGATCAACACAGTTTCTATTTGGGCTGTATTAAGATGA 1200
QY 2565 GGGGAGAGAGATGAGAGACATCTTCACTGCGGCTGTCTCAGGGGAGAGACCTGAT 2624
DB 1201 GGGGAGAGAGATGAGAGACATCTTCACTGCGGCTGTCTCAGGGGAGAGACCTGAT 1260
QY 2625 GCAAGTGGCTCCGCTCCGAGAGCTGCTGAGAGCAACCTCTGAGAGTCCCTCACTG 2684
DB 1261 GCAAGTGGCTCCGCTCCGAGAGCTGCTGAGAGCAACCTCTGAGAGTCCCTCACTG 1320
QY 2685 CTTGTACAGAGACTCGAGAACAAACCTTCTGACACAGTATGAGGCCATTTGAGAGAAAT 2744
DB 1321 CTTGTACAGAGACTCGAGAACAAACCTTCTGACACAGTATGAGGCCATTTGAGAGAAAT 1380
QY 2745 GGGCATGTGTGTGAGAAACAGACATGAGAGCTTTAGTGTGCACTTTCTGATTAATTTAG 2804
DB 1381 GGGCATGTGTGTGAGAAACAGACATGAGAGCTTTAGTGTGCACTTTCTGATTAATTTAG 1440
QY 2805 CCGCACGTAAGAGAGCTTCACTGATTTGAGGGGAGGAGCAGACAGATCAACATGAGAGCC 2864
```


Db	1441	CCGCCAGTGAAGAAGCTTACGCTGATTGAGGGCAGGGCAGACGATCAAGATCAATGAGACC	1500
Qy	2865	CACCATGGTAGTCCGTGTTCAAGTGGGATCCCAAGTCAGATGCTATTGGCAGATTCTT	2924
Db	1501	CACCATGGTAGTCCGTGTTCAAGTGGGATCCCAAGTCAGATGCTATTGGCAGATTCTT	1560
Qy	2925	CTCCGTCTCTAAGGTCACAGAAACCTGGAAGAGCTGACCTTAAGTGAACCTGCTGAG	2984
Db	1561	CTCCGTCTCTAAGGTCACAGAAACCTGGAAGAGCTGACCTTAAGTGAACCTGCTGAG	1620
Qy	2985	CCACTGTCAGTGAAGAGTCTTTTAAAGACCTGAGAGGCTTCGCTGCTCCGAGAC	3044
Db	1621	CCACTGTCAGTGAAGAGTCTTTTAAAGACCTTGAAGGCTTCGCTGCTCCGAGAC	1680
Qy	3045	CTTCGGTGTGGCTGGCTGTGGCTCTCAAGCTGAGAGCTGCAAGGACCTTTCGAGCT	3104
Db	1681	CTTCGGTGTGGCTGTGGCTCTCAAGCTGAGAGCTGCAAGGACCTTTCGAGCT	1740
Qy	3105	GAGAGCCAACACAGACCTTCGACCGAGCTTGAGCTTGAGCTTCAATGTGCTCAAGATGCTGG	3164
Db	1741	GAGAGCCAACACAGACCTTCGACCGAGCTTGAGCTTCAATGTGCTCAAGATGCTGG	1800
Qy	3165	AGCCAAACACCTTTGCCAGAGACTGABAACAGCCGAGCTGCAAGCTTAAGAGACCTGACGT	3224
Db	1801	AGCCAAACACCTTTGCCAGAGACTGABAACAGCCGAGCTGCAAGCTTAAGAGACCTGACGT	1860
Qy	3225	GGTCAGCTGTGGCTCAAGTCTGACTGCTGACAGACCTTGACCTGTGCTTAGTGCAG	3284
Db	1861	GGTCAGCTGTGGCTCAAGTCTGACTGCTGACAGACCTTGACCTGTGCTTAGTGCAG	1920
Qy	3285	CCCCAGCTTGAAGAGCTTGAAGCTTCGACAGACCAACCTGATGACGTTGGCGTGCAGT	3344
Db	1921	CCCCAGCTTGAAGAGCTTGAAGCTTCGACAGACCAACCTGATGACGTTGGCGTGCAGT	1980
Qy	3345	GCCTGTGGAAGGGGCTCAGGCAATCCTGCTGCAACTCATACGCTGGGGCTGACCAAGC	3404
Db	1981	GCCTGTGGAAGGGGCTCAGGCAATCCTGCTGCAACTCATACGCTGGGGCTGACCAAGC	2040
Qy	3405	AACTCTAGTGAATGAGATGAGGCAAGAACTGAGGGCCTGGAGCAGAGAAACCTCAGCT	3464
Db	2041	AACTCTAGTGAATGAGATGAGGCAAGAACTGAGGGCCTGGAGCAGAGAAACCTCAGCT	2100
Qy	3465	GCTCATCTTCAAGCAGACGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGATACGGG	3524
Db	2101	GCTCATCTTCAAGCAGACGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGATACGGG	2160
Qy	3525	AGAGATGATTAATGCAATCTCTCACTCAGCGGCAGAGACTCCGATTCAGAGAGGCGGC	3584
Db	2161	AGAGATGATTAATGCAATCTCTCACTCAGCGGCAGAGACTCCGATTCAGAGAGGCGGC	2220
Qy	3585	TTCCATGTTGCTCAGGCTAATCTCAAACTCTGGAAGTGAAGAAATCTTCCAAATTCG	3644
Db	2221	TTCCATGTTGCTCAGGCTAATCTCAAACTCTGGAAGTGAAGAAATCTTCCAAATTCG	2280
Qy	3645	TGAGATTGCAAGAGAAAGCTCCCAAGAGTATGATCCGGTGAACCTTGTGTGCGCTTC	3704
Db	2281	TGAGATTGCAAGAGAAAGCTCCCAAGAGTATGATCCGGTGAACCTTGTGTGCGCTTC	2340
Qy	3705	TTCGTGCTCTCAAGGGGACCTGCAATAGAGGCTTTGGGGACTGACGATGACTTCTGGGG	3764
Db	2341	TTCGTGCTCTCAAGGGGACCTGCAATAGAGGCTTTGGGGACTGACGATGACTTCTGGGG	2400
Qy	3765	CCCCACGGGGCTGTGGCTACTGAGAGTATGACAAAGAAAGAACTTGTACGAGTTCA	3824
Db	2401	CCCCACGGGGCTGTGGCTACTGAGAGTATGACAAAGAAAGAACTTGTACGAGTTCA	2460
Qy	3825	CTTCCCTGTAGCTGCTCTTACCCGCTGGCCCAACAGGGCTCTGCTTGTGTATGAGAGA	3884
Db	2461	CTTCCCTGTAGCTGCTCTTACCCGCTGGCCCAACAGGGCTCTGCTTGTGTATGAGAGA	2520
Qy	3885	AGCGGTGACCTTGAATTAATTTCTGTGTGGAACAATTCTGGGTGAGATCAACC	3944

D	2521	AGGGGTGACCGCTTGAGATTGAATTCGTGTGTGTGGGACCAAGTTCCTGGGTAGATCAACCC	2580
Q	3945	ACAGCACAGCTGGATGGTGGCAGAGGCGCTCTGTGGACATCAAGACCTGAGCTGAGCTGT	4004
D	2581	ACAGCACAGCTGGATGGTGGCAGAGGCGCTCTGTGGACATCAAGACCTGAGCTGAGCTGT	2640
Q	4005	GGAGCTGTACCTCCCTCCACTTTTGTGGCTCTCCCAAGGGGGCCATGTGACATCTCCCT	4064
D	2641	GGAAGCTGTACCTCCCTCCACTTTTGTGGCTCTCCCAAGGGGGCCATGTGACATCTCCCT	2700
Q	4065	GTTTCCAATGGGCCCACTTTAAAGAGAGGGGATGCTCCTGGAAAGCAGCCAGGGTGA	4124
D	2701	GTTTCCAAGTGGGCCCACTTTAAAGAGAGGGGATGCTCCTGGAAAGCAGCCAGGGTGA	2760
Q	4125	GCTGCATCACATAGTTCGTGAAAAACCCAGCTTCTCCCTTGGAGTCTCTGTAAAT	4184
D	2761	GCTGCATCACATAGTTCGTGAAAAACCCAGCTTCTCCCTTGGAGTCTCTGTAAAT	2820
Q	4185	GATCCATTAATGCCCTGGCTTCAATTCCCGTCACTCTGTGTGTGTGTTTACACCGCGT	4244
D	2821	GATCCATTAATGCCCTGGCTTCAATTCCCGTCACTCTGTGTGTGTGTTTACACCGCGT	2880
Q	4245	CCATCCGAGGAAGTCACTTCCACCGCTTACCGATGCCAAGGACGTGCTCAATTGG--	4302
D	2881	CCATCCGAGGAAGTCACTTCCACCGCTTACCGATGCCAAGGACGTGCTCAATTGGAA	2940
Q	4303	-----	4302
D	2941	GGCCATGATGATCTTGAAATGAATTCGAGTTTGTGTGCAATCCACAGCCACCCCGCT	3000
Q	4303	-----	4302
D	3001	GACCCCACTTATATGAGGCTGTGCTTACACTGTCTGGTCTGTTCAGGAGTGTGGA	3060
Q	4303	-----AAGAACTGGAGCTGTGCTATCGAAGCCCTGGAGAAAGACAGCGTTTCTC	4352
D	3061	AATATCTCCCAAGGAATGAGTCTGTCTATCGAAGCCCTGGAGAAAGACAGCGTTTCTC	3120
Q	4353	GGAGTTTACGTTTGGCCACTTGGGATCAGGGATCAGGCTTGCAAGTGAAGACAAGAAAG	4412
D	3121	GGAGTTTACGTTTGGCCACTTGGGATCAGGGATCAGGCTTGCAAGTGAAGACAAGAAAG	3180
Q	4413	TGAGACTCTGTGTGTGGAGGCGCTTGGTGAACCAAGAGATCTCATGCTGCAACTACTCT	4472
D	3181	TGAGACTCTGTGTGTGGAGGCGCTTGGTGAACCAAGAGATCTCATGCTGCAACTACTCT	3240
Q	4473	GATCCCTCCAGCCCGCATAGCCGTAACCTTCAGTGAAGCCCGCAGTGTGTGACATT	4532
D	3241	GATCCCTCCAGCCCGCATAGCCGTAACCTTCAGTGAAGCCCGCAGTGTGTGACATT	3300
Q	4533	TGTGAGACCAATATCGAAGACAGCTGATAGCCCGAGTACATCGGTGAGAGGTTTCTTGA	4592
D	3301	TGTGAGACCAATATCGAAGACAGCTGATAGCCCGAGTACATCGGTGAGAGGTTTCTTGA	3360
Q	4593	CAAACTGCATGACAGAGTCTGAGCCAGCAAGACAGTACGAGAGGGTGTGCTAGAAAC	4652
D	3361	CAAACTGCATGACAGAGTCTGAGCCAGCAAGACAGTACGAGAGGGTGTGCTAGAAAC	3420
Q	4653	GAGGCCACGCAATATGGGAAAGCTGTTCAGCTTGAACAGTACTCTGGGACCGGAAGTGA	4712
D	3421	GAGGCCACGCAATATGGGAAAGCTGTTCAGCTTGAACAGTACTCTGGGACCGGAAGTGA	3480
Q	4713	AGATGACACTTACCAAGCCCTGAAAGAGACCCATCTCTCATTTATGAACTTGGGA	4772
D	3481	AGATGACACTTACCAAGCCCTGAAAGAGACCCATCTCTCATTTATGAACTTGGGA	3540
Q	4773	GAAAGGACGCAAAAAGGAGCTCTGTGCCACTCAGACGCTGAAGTATCAACACAGCCCTTG	4832
D	3541	GAAAGGACGCAAAAAGGAGCTCTGTGCCACTCAGACGCTGAAGTATCAACACAGCCCTTG	3600
Q	4833	ACCCTTGAATGTCGTGGCTTGTGGAGCAACCTTCTTGTGGTCTCAATTTCTTCTGCAAA	4892
D	3601	ACCCTTGAATGTCGTGGCTTGTGGAGCAACCTTCTTGTGGTCTCAATTTCTTCTGCAAA	3660

QY 4893 AAGTGCATCTGTTGCTTCCAGACCTAAGTAATGAATCTTGATGATGCTTTG 4952
 DB 3661 AAGTGCATCTGTTGCTTCCAGACCTAAGTAATGAATCTTGATGATGCTTTG 3720
 QY 4953 TGGGCAATATGTTCCATGCGCAGGAGTGCACAGGGGGCCCGCAGTCCAGTGGCTTAA 5012
 DB 3721 TGGGCAATATGTTCCATGCGCAGGAGTGCACAGGGGGCCCGCAGTCCAGTGGCTTAA 3780
 QY 5013 GCATCTCAGGAAATGTTCCATCTGAGCTGGCAAGACCCCTGACACCTCAATAGAGCTCA 5072
 DB 3781 GCATCTCAGGAAATGTTCCATCTGAGCTGGCAAGACCCCTGACACCTCAATAGAGCTCA 3840
 QY 5073 TCTGTTGCGCCACAGCGCCAGCTAGAGCTTCCGATTCCTCCATCCAGGGGCAAGAGGA 5132
 DB 3841 TCTGTTGCGCCACAGCGCCAGCTAGAGCTTCCGATTCCTCCATCCAGGGGCAAGAGGA 3900
 QY 5133 ATAGAGAGGACATGAGACCATTTGCTGCTGTGTGTACAGGGGTGAGCCCCAAATTTG 5192
 DB 3901 ATAGAGAGGACATGAGACCATTTGCTGCTGTGTGTACAGGGGTGAGCCCCAAATTTG 3960
 QY 5193 GATTCAAGCTGGAGGCGCAGGTGATTTCTTGCTTTGTAACAGGAATCTTACAGAGCA 5252
 DB 3961 GATTCAAGCTGGAGGCGCAGGTGATTTCTTGCTTTGTAACAGGAATCTTACAGAGCA 4020
 QY 5253 GCCAAGAGTAAAGTGAAGAGTTTATTCAGAAATTAAGAGTATCAAGCTCTT 5312
 DB 4021 GCCAAGAGTAAAGTGAAGAGTTTATTCAGAAATTAAGAGTATCAAGCTCTT 4080
 QY 5313 TAGAA 5317
 DB 4081 TAAAA 4085

RESULT 15
 AX089769 4556 bp DNA linear PAT 21-MAR-2001
 LOCUS Sequence 9 from Patent WO0116170.
 DEFINITION AX089769
 ACCESSION AX089769
 VERSION AX089769.1 GI:13443941
 KEYWORDS
 SOURCE
 ORGANISM
 .
 synthetic construct
 synthetic construct
 artificial sequences.
 REFERENCE
 1. Reed, J.C.
 AUTHORS Card proteins involved in cell death regulation
 TITLE Patent: WO 0116170-A 9 08-MAR-2001;
 JOURNAL The Burnham Institute (US)
 FEATURES
 source
 1. 4556
 Location/Qualifiers
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 1. 4365
 /note="unnamed protein product"
 /codon_start=1
 /translation="1"
 /protein_id="CAC34903.1"
 /db_xref="GI:13443942"
 /translation="MAGGAMGRVLAFLYELKKKEELKEPOLLANKHSHSSSGEPAP
 PKTSGMVAAYLVIAVQGBORAMDIALHMEQGLSLCAQOEBGAGHSPSPYSSSE
 PHIGSPQSTSTAVLMPVTHEIPAGCTQSSERVLRQLDPISGRKRELSASULTAL
 PSSPDHESPSQSPNAPSTAVGVSGSPQPSLHAREAGETQPLDETSTGIYYTE
 IBERERESKEKSPMAAVVGTPOAHTLQPHHWEPSVESLSTPMKNEEDFNO
 KFTQILLORPHRSDPLVKRSMPTVEENRGLLEIDLFPGIDTQEPRIVLIOG
 AAGIGKSTLAVQKEMRGOXYGDFHOFYFSCRELAOSKVSIAELIGDGTAP
 APIROILSRPERLLFLIDVDRGWLQSPSEELCIHMSQOPADALISLAKTILP
 EASFLTATTATLQNLIPSLBOARWVYLVGSESSKEKFFRYITDERPAITAFRLV
 SNEKMAALCLVPSWNLACTCLMQKREKELTISKTTTCLHLYLAQALDAQPLG
 QLRDLCSIAAGVIMQKTLFSPDLRKHLDAIISTFLKMGILQHPRLPDAYSFHL
 CROEPMAASVYLEDKGRGKSNCCIIDEKTLFAGIGLFGASTTRFLGLISDEG
 EREMENTFHCRLSGRNLNQWVPSLQILLQPSHLSBLHCLVETRNKTFVQVMAHEE

CDS
 /note="unnamed protein product"
 /codon_start=1
 /translation="1"
 /protein_id="CAC34903.1"
 /db_xref="GI:13443942"
 /translation="MAGGAMGRVLAFLYELKKKEELKEPOLLANKHSHSSSGEPAP
 PKTSGMVAAYLVIAVQGBORAMDIALHMEQGLSLCAQOEBGAGHSPSPYSSSE
 PHIGSPQSTSTAVLMPVTHEIPAGCTQSSERVLRQLDPISGRKRELSASULTAL
 PSSPDHESPSQSPNAPSTAVGVSGSPQPSLHAREAGETQPLDETSTGIYYTE
 IBERERESKEKSPMAAVVGTPOAHTLQPHHWEPSVESLSTPMKNEEDFNO
 KFTQILLORPHRSDPLVKRSMPTVEENRGLLEIDLFPGIDTQEPRIVLIOG
 AAGIGKSTLAVQKEMRGOXYGDFHOFYFSCRELAOSKVSIAELIGDGTAP
 APIROILSRPERLLFLIDVDRGWLQSPSEELCIHMSQOPADALISLAKTILP
 EASFLTATTATLQNLIPSLBOARWVYLVGSESSKEKFFRYITDERPAITAFRLV
 SNEKMAALCLVPSWNLACTCLMQKREKELTISKTTTCLHLYLAQALDAQPLG
 QLRDLCSIAAGVIMQKTLFSPDLRKHLDAIISTFLKMGILQHPRLPDAYSFHL
 CROEPMAASVYLEDKGRGKSNCCIIDEKTLFAGIGLFGASTTRFLGLISDEG
 EREMENTFHCRLSGRNLNQWVPSLQILLQPSHLSBLHCLVETRNKTFVQVMAHEE

BASE COUNT 1100 a 1223 c 1256 g 977 t
 ORIGIN
 MCMCVETDMEILLVCTFCTKESHVKKLLQILIBORHRSWTSPVTVVLFVWVPTDAWQ
 ILFSLVKTATRLKELDLISGNSLSHAYVSLCTTIRPRCLLETLLVAGCGVLAEDCKD
 LAFGLRANQVLTLELDLFRVLTDAQAKHLCDLRPSCKLQRLVLSCLTSDCCDL
 ASVLASPSLKELDLQNNLDVGRVLCBGRHPACKLIRGLQVTLISDMROELK
 ALEOKPOLILFRRKPSVMTPEGLDGENENSTSLKRLGRSLGSEAAASHOANLK
 LLDVKRIPILAIRESSEVYVAVGLDVPASAGODLTKGLDGDDELEGANDV
 ELIDSTNRYSVMPPTAGVYVMSATGLDFIAPDEYVTVTAFGSMOHLADIQHEOW
 LVGGLFVDTAARPEBAVAHILPHRISLQGEVDVSVFLVAHFRKNEQVLEHARVEP
 YAVLESFSLMGLILRLIASGTRLSIPITSNLTLYVHPHEDIKHLIYVPSDALITK
 AIDDEDRFHGVRQLQTSPPMBELNVSIVSNRNLKVMPELKSYPSPGELIHF
 KEYAGMEKPEIQLETEKRHGTLVMDTEVKKPDLQLVAAAPPSGAFAVENRHL
 QARMGDLKGVLDLQDNEVLTENEKELVBEKTRQSKNEALLSWEKKGDLALDVLFR
 SISERDPLVSVLRQONL"
 Query Match 62.5%; Score 3400.4; DB 6; Length 4556;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 3555; Conservative 0; Mismatches 236; Indels 3; Gaps 1;
 QY 523 ATGGCTGGCGGAGCCTGGGGCGGCTGGCTGTGTAATTGGAAGTTCTTGAAGAGAGAG 582
 DB 1 ATGGCTGGCGGAGCCTGGGGCGGCTGGCTGTGTAATTGGAAGTTCTTGAAGAGAGAG 60
 QY 583 CTGAAGAGATTCCAGCTTGTGCTGCGCAATTAAGGCACTCCAGAGCTCTTCGGGTGAG 642
 DB 61 CTGAAGAGATTCCAGCTTGTGCTGCGCAATTAAGGCACTCCAGAGCTCTTCGGGTGAG 120
 QY 643 ACACCGCTCAGCCAGAGAGAGAGAGTGCATGAGAGTGGCTCGTACCTGGTGGCTCAG 702
 DB 121 ACACCGCTCAGCCAGAGAGAGAGAGTGCATGAGAGTGGCTCGTACCTGGTGGCTCAG 180
 QY 703 TATGGGAGACAGCGGCTGGGACCTAGCCTTCATACCTGGGAGACAGATGGGGCTGAG 762
 DB 181 TATGGGAGACAGCGGCTGGGACCTAGCCTTCATACCTGGGAGACAGATGGGGCTGAG 240
 QY 763 TCACGTGGGCGCCAGAGCCAGAGAGAGGAGAGGAGCCACTCCCTCAATCCCTCAAGCCCA 822
 DB 241 TCACGTGGGCGCCAGAGCCAGAGAGAGGAGAGGAGCCACTCCCTCAATCCCTCAAGCCCA 300
 QY 823 AGTGAACCCCACTGGGGCTCTCCAGCCCAACCACTCCACCGCAGTGAATGGCTTGG 882
 DB 301 AGTGAACCCCACTGGGGCTCTCCAGCCCAACCACTCCACCGCAGTGAATGGCTTGG 360
 QY 883 ATTCATGAATTCGCGGCGGGGTGCAACCAAGGCTCAGAGAGAGAGGATTTGAGACCTG 942
 DB 361 ATTCATGAATTCGCGGCGGGGTGCAACCAAGGCTCAGAGAGAGAGGATTTGAGACCTG 420
 QY 943 CCTGACATCTGAGAGCGCGCTGGAGAGAAATCTGCTCACTGCTCAACCAAGCTTT 1002
 DB 421 CCTGACATCTGAGAGCGCGCTGGAGAGAAATCTGCTCACTGCTCAACCAAGCTTT 480
 QY 1003 CCAGCTCCCGAGACCATGATGCTCCAGCCAGAGATCAACCAAGCCCGCCCATCTCA 1062
 DB 481 CCAGCTCCCGAGACCATGATGCTCCAGCCAGAGATCAACCAAGCCCGCCCATCTCA 540
 QY 1063 GCAAGTCTGGGGAGCTGGGGATTCCTCAAGCTCAGCCCAAGCTTACAGCCAGAGAGAG 1122
 DB 541 GCAAGTCTGGGGAGCTGGGGATTCCTCAAGCTCAGCCCAAGCTTACAGCCAGAGAGAG 600
 QY 1123 GTCCTGGGAGACCATGAGCTCTGATGAAGAGTCAAGAAATTTATCAACAGAAATTCAG 1182
 DB 601 GTCCTGGGAGACCATGAGCTCTGATGAAGAGTCAAGAAATTTATCAACAGAAATTCAG 660
 QY 1183 GAAAGAGAGAGAGAAATTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1242
 DB 661 GAAAGAGAGAGAGAAATTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 1243 CCCCCAGGCGCACAACAGCTTACAGCCCAACCACTGGGAGGCTTCTGTGAGA 1302
 DB 721 CCCCCAGGCGCACAACAGCTTACAGCCCAACCACTGGGAGGCTTCTGTGAGA 780
 QY 1303 GAGAGCTCTGTTCACATGGCCCTGAGAAATTAAGAGATTTTAAACAAATTCACACAG 1362


```
Db 781 GAGAGCTCTGTTCCATGAGCCCTGGAAAAATAGATTTTAAACAAAATTACACAG 840
Qy 1363 CTGCTACTTTCTACAAAGACTCTCAACCCAGAAAGCCAGATCCCTGTCTCAAGAGAGCTGG 1422
Db 841 CTGCTACTTTCTACAAAGACTCTCAACCCAGAAAGCCAGATCCCTGTCTCAAGAGAGCTGG 900
Qy 1423 CCGATTATATGAGGAGAGAAATCGAGGACATTTATTTAGATTCAGAGACTTATTTGGCCCA 1482
Db 901 CCGATTATATGAGGAGAGAAATCGAGGACATTTATTTAGATTCAGAGACTTATTTGGCCCA 960
Qy 1483 GGCCTGATATCCCAAGAACTCCGATATGATCTATGAGGGGGGTGTCTGTAATTGGGAAG 1542
Db 961 GGCCTGATATCCCAAGAACTCCGATATGATCTATGAGGGGGGTGTCTGTAATTGGGAAG 1020
Qy 1543 TCAACACTGGCCAGGCGAGGTGAAGAAAGCTGGGGGAGAGGCGAGCTGTATGGGGACCG 1602
Db 1021 TCAACACTGGCCAGGCGAGGTGAAGAAAGCTGGGGGAGAGGCGAGCTGTATGGGGACCG 1080
Qy 1603 TTCCAGATATGTTCTTCTCTCAAGCTGAGAGAGCTGGCCCAAGTCCAGAGTGTAGTCTC 1662
Db 1081 TTCCAGATATGTTCTTCTCTCAAGCTGAGAGAGCTGGCCCAAGTGTAGTAGTCTC 1140
Qy 1663 GCTGAGCTCATCGGAAAGATGGAGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAGATGGAGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1200
Qy 1723 AGGCGAGAGCGGCTGCTCTTCTATCTCGATGAGTGTAGATGAGCCAGAGATGGGCTTTCAG 1782
Db 1201 AGGCGAGAGCGGCTGCTCTTCTATCTCGATGAGTGTAGATGAGCCAGAGATGGGCTTTCAG 1260
Qy 1783 GAGCGGAGTTCTGAGCTCTGTCTGCACTGAGACAGCCAGCGGGGGAGATGACATGCTG 1842
Db 1261 GAGCGGAGTTCTGAGCTCTGTCTGCACTGAGACAGCCAGCGGGGGAGATGACATGCTG 1320
Qy 1843 GGCAGTTTGTCTGGGAGAAACTATACTTCCGAGGACATCTTCTGATCAAGGCTCGGAGC 1902
Db 1321 GGCAGTTTGTCTGGGAGAAACTATACTTCCGAGGACATCTTCTGATCAAGGCTCGGAGC 1380
Qy 1903 ACAGCTCTGCAGAACTCTCATCTCTTTGGAGCAGGACAGTTGGGTAGAGTCTGAGG 1962
Db 1381 ACAGCTCTGCAGAACTCTCATCTCTTTGGAGCAGGACAGTTGGGTAGAGTCTGAGG 1440
Qy 1963 TTCTCTAGTCCAGAGAGAGAAATTTTCTACAGATATTTTCAAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTAGTCCAGAGAGAGAAATTTTCTACAGATATTTTCAAGATGAAAGGCAAGCA 1500
Qy 2023 ATTAGAGCTTTAGGTTGGTCAATCAAAAGAGCTTGAGGCTCTGTCTTGTGCC 2082
Db 1501 ATTAGAGCTTTAGGTTGGTCAATCAAAAGAGCTTGAGGCTCTGTCTTGTGCC 1560
Qy 2083 TGGGTGTCTGGCTGGCTGACCTTGCTGTGATGAGCAGATGAAAGGAAAGAAAATC 2142
Db 1561 TGGGTGTCTGGCTGGCTGACCTTGCTGTGATGAGCAGATGAAAGGAAAGAAAATC 1620
Qy 2143 ACACTGACTTTCAGAGACCAACCAACCTCTGTCTACATTAACCTTGCCAGGCTCTCAA 2202
Db 1621 ACACTGACTTTCAGAGACCAACCAACCTCTGTCTACATTAACCTTGCCAGGCTCTCAA 1680
Qy 2203 GCTCAGCATTGGAGCCCAAGCTCAGAGACTCTGTCTCTGGGCTGTGAGGGGATCTGG 2262
Db 1681 GCTCAGCATTGGAGCCCAAGCTCAGAGACTCTGTCTCTGGGCTGTGAGGGGATCTGG 1740
Qy 2263 CAAAAAAGACCTTTTCACTCAGATGACCTCAGAGAGCAATGGGTTAGATGGGCAATC 2322
Db 1741 CAAAAAAGACCTTTTCACTCAGATGACCTCAGAGAGCAATGGGTTAGATGGGCAATC 1800
Qy 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCAACCCATCTCTTGAAGTACAGC 2382
Db 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCAACCCATCTCTTGAAGTACAGC 1860
Qy 2383 TTCAATTACCTCTGTTTCAAGAGTCTTTCAGAGATGTCTTATGTCTTGGAGATGAG 2442
Db 1861 TTCAATTACCTCTGTTTCAAGAGTCTTTCAGAGATGTCTTATGTCTTGGAGATGAG 1920
Qy 2443 AAGGGGAGAGTAAACAATTCTAATTGCATATGATTTGGAAAGACGCTAAGACATAT 2502
Db 1921 AAGGGGAGAGTAAACAATTCTAATTGCATATGATTTGGAAAGACGCTAAGACATAT 1980
Qy 2503 GGAATTCATGCGCTGTTTGGGGGATCAACCAAGTTTCTATTTGGGCTGTTAAGAT 2562
Db 1981 GGAATTCATGCGCTGTTTGGGGGATCAACCAAGTTTCTATTTGGGCTGTTAAGAT 2040
Qy 2563 GAGGGGAGAGAGATGAGAAACATCTTTCAGTCCGAGCTGCTCAGGGGAGAACTG 2622
Db 2041 GAGGGGAGAGAGATGAGAAACATCTTTCAGTCCGAGCTGCTCAGGGGAGAACTG 2100
Qy 2623 ATGCAATGGGTCCCGTCCCTGACAGTGTCTGAGCCACATCTCTGAGTCCCTCAC 2682
Db 2101 ATGCAATGGGTCCCGTCCCTGACAGTGTCTGAGCCACATCTCTGAGTCCCTCAC 2160
Qy 2683 TGGCTGTAGAGACTGGGAAACAAAAGTTCTGACACAATGATGGGCCATTTTCAGAA 2742
Db 2161 TGGCTGTAGAGACTGGGAAACAAAAGTTCTGACACAATGATGGGCCATTTTCAGAA 2220
Qy 2743 ATGGGCAATGTGTAGAAACAGACATGAGACTTTAGTGTGCACTTTCGATTAATTC 2802
Db 2221 ATGGGCAATGTGTAGAAACAGACATGAGACTTTAGTGTGCACTTTCGATTAATTC 2280
Qy 2803 AGCTGCCAGTGAAGAAAGCTTCACTGATTTAGGGGAGGAGACAGATCAATGAGAGC 2862
Db 2281 AGCTGCCAGTGAAGAAAGCTTCACTGATTTAGGGGAGGAGACAGATCAATGAGAGC 2340
Qy 2863 CCGACCATGGTATTCCTGTTCAAGTGGTCCAGTGCACANATCCCTATTTGGCAGATTCTC 2922
Db 2341 CCGACCATGGTATTCCTGTTCAAGTGGTCCAGTGCACANATCCCTATTTGGCAGATTCTC 2400
Qy 2923 TTCTCCGTCTCAAGGCTCACAGAAACCTGAGAGAGCTGAGACCTTAAGTGAATTCGCTG 2982
Db 2401 TTCTCCGTCTCAAGGCTCACAGAAACCTGAGAGAGCTGAGACCTTAAGTGAATTCGCTG 2460
Qy 2983 AGCCACTCTGAGTGAAGATCTTTTGTAGACCTTGAAGACGCTCTGCTGCTCTGAG 3042
Db 2461 AGCCACTCTGAGTGAAGATCTTTTGTAGACCTTGAAGACGCTCTGCTGCTCTGAG 2520
Qy 3043 ACCCTGCGTTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3102
Db 2521 ACCCTGCGTTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 3103 CTGAGAGCCAAACAGACCTTGAACCGAGCTGAGCTTCAATGTGCTCAAGGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTTGAACCGAGCTGAGCTTCAATGTGCTCAAGGATGCT 2640
Qy 3163 GAGAGCCAAACAGCTTTCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGAGCTGAG 3222
Db 2641 GAGAGCCAAACAGCTTTCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGAGCTGAG 2700
Qy 3223 CTGCTCAGCTGTGGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
Db 2701 CTGCTCAGCTGTGGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Qy 3283 AGCCCGAGCTGAAAGAGCTAGACCTGCAACAGAAACAATCTGATGAGCTTGGGCTGCA 3342
Db 2761 AGCCCGAGCTGAAAGAGCTAGACCTGCAACAGAAACAATCTGATGAGCTTGGGCTGCA 2820
Qy 3343 CTGCTCTGTAAGGGGCTCAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3402
Db 2821 CTGCTCTGTAAGGGGCTCAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
Qy 3403 ACAACTCTGATGATATGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3462
Db 2881 ACAACTCTGATGATATGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
Qy 3463 CTGCTCATCTTTCAGAGACGGAACCAAGTGTATGACCCCTTACTGAGGGCTGAGATAGC 3522
Db 2941 CTGCTCATCTTTCAGAGACGGAACCAAGTGTATGACCCCTTACTGAGGGCTGAGATAGC 3000
```

